Б.П. ЛАТХИ

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Б. П. Латхи

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

COMMUNICATION SYSTEMS

B. P. LATHI

Associate Professor of Electrical Engineering
Bradley University

John Wiley & Sons, Inc.

New York · London · Sydney

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Б. П. ЛАТХИ

Перевод с английского
под общей редакцией
Б. И. Кувшинова

Допущено
Министерством связи СССР
в качестве учебного пособия
для электротехнических институтов
связи



Издательство «Связь»
Москва 1971

6Ф2
Л-27

УДК 621.391.3

Латхи Б. П.

Системы передачи информации.

Л-27

Перевод с англ. под общей редакцией Б. И. Кувшинова. М., «Связь», 1971, 324 с. с илл. Библ. 20

В книге излагаются основы теории сигналов, систем передачи информации, вопросы помехоустойчивости систем связи, а также основы теории информации.

Отличительной особенностью книги является простота изложения и физическое объяснение основных понятий и теоретических положений. Изучение материала книги облегчают многочисленные примеры и иллюстрации. В конце каждой главы приводятся задачи.

Книга может служить учебным пособием студентам электрорадиотехнических специальностей, а также будет полезна инженерам-практикам и широкому кругу специалистов, интересующихся теорией передачи информации.

6Ф2

3—6—1
6—71

Б. П. Латхи

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Редакторы *Н. К. Логинова, Г. Ш. Погорельская* Художник *Г. Р. Левин*
Техн. редактор *К. Г. Маркоч* Корректор *Л. Н. Лещева*

Сдано в набор 18/II 1971 г. Подписано в печ. 25/V 1971 г.
Форм. бум. 60×90/16 20,0 печ. л. 20,0 усл.-п. л. 20,17 уч.-изд. л.
Бумага книжно-журнальная. Тираж 20 000 экз. Цена 95 коп. Зак. изд. 14803
Издательство «Связь», Москва-центр, Чистопрудный бульвар, 2

Типография издательства «Связь» Комитета по печати при Совете Министров СССР. Москва-центр, ул. Кирова, 40. Зак. тип. 41

От издательства

Предлагаемая вниманию читателей книга Б. П. Латхи «Системы передачи информации» посвящена некоторым важным вопросам теории передачи сигналов: анализу сигналов, теории непрерывной и дискретной модуляции, теории помехоустойчивости систем передачи, основам теории информации. В настоящее время по этим вопросам имеется большое число монографий, учебников, статей в периодической печати. Однако в большинстве случаев чтение такой литературы представляет большие трудности для студентов вузов, особенно младших курсов, как вследствие недостаточной подготовки по теории вероятностей и теории случайных процессов, так и вследствие формально-математического характера изложения материала, принятого многими авторами.

Книга Б. П. Латхи составляет в этом смысле приятное исключение. Она рассчитана на студентов электрорадиотехнических специальностей, прослушавших курс «Теория линейных цепей» и желающих расширить свои знания в области теории систем передачи информации. В ней удачно сочетаются математическая строгость изложения и физическое толкование результатов, что позволяет донести содержание книги даже до неподготовленного читателя и значительно облегчает понимание таких тонких и сложных вопросов современной теории передачи сигналов, как помехоустойчивость систем связи, методы оптимального приема сигналов, обмен полосы частот на отношение сигнал/шум и др. Книга снабжена большим числом примеров и иллюстраций. В конце каждой главы приводятся задачи, решение которых позволяет читателю контролировать степень усвоения материала. Для получения необходимых результатов автор обходится минимальным числом статистических характеристик случайных процессов: спектром плотности мощности и функцией плотности вероятности, разъяснение которых приводится в книге и не требует привлечения других источников. Можно надеяться, что книга Б. П. Латхи позволит начинающему читателю составить представление о современных системах передачи информации и облегчит их дальнейшее более глубокое и серьезное изучение, основанное на использовании аппарата теории случайных процессов.

Автор широко использует спектральный анализ детерминированных и случайных сигналов, основы которого рассматриваются в первых двух главах книги. Следует отметить также использование теорем частотной и временной свертки при нахождении спектров многих сигналов, что является весьма удачным методическим приемом.

В трех последующих главах рассматриваются системы связи с различными видами модуляции, где наряду с «традиционными» вопросами обсуждается и ряд других: искажения при детектировании сигналов, понятие линейной и нелинейной модуляции, системы с частично-подавленной боковой и др. В шестой главе рассматриваются шумы элементов и цепей радиоаппаратуры.

В остальных трех главах изучаются помехоустойчивость систем связи, элементы теории информации и системы передачи дискретной информации.

Перевод книги выполнен кандидатами технических наук Б. И. Кувшиновым, М. В. Назаровым и С. С. Свириденко практически без сокращений; исправлены лишь замеченные опечатки в формулах и даны примечания в тех местах книги, которые этого требовали. Литература, не известная советскому читателю, дополнена соответствующими отечественными источниками.

Отзывы и замечания по книге следует направлять в издательство «Связь» (Москва-центр, Чистопрудный бульвар, 2).

Предисловие

Цель этой книги — дать студентам предварительные сведения о системах передачи информации и общих принципах современной теории передачи информации на ранней стадии обучения. Книга начинается с рассмотрения характерных систем передачи информации, а затем раскрывается роль отношения сигнал/шум и полосы частот в ограничении скорости передачи информации.

Поскольку данная книга предназначена в качестве вводного курса, в ней не затронуты многие специальные вопросы, касающиеся спектров плотности мощности случайных процессов. Студенты знакомятся со спектром плотности мощности неслучайных сигналов, а затем без каких-либо формальных выводов это понятие переносится на случайные сигналы. По-видимому, строгое рассмотрение случайных процессов в таком вводном курсе отвлекло бы внимание читателя и противоречило бы самому назначению курса.

Основное внимание в книге уделяется физическому толкованию понятий, а не математическим выкладкам. В этом отношении она близка к моей более ранней книге «Сигналы, системы и передача информации». Основные понятия теории информации вводятся не как аксиомы, а на основе эвристических рассуждений.

Книга рассчитана на изучение в течение одного семестра или квартала при соответствующем выборе тематики. Можно рекомендовать несколько целесообразных последовательностей изучения глав книги с обязательным изучением первых четырех глав, а затем гл. 7 и 8, или 6, 7 и 8, или 7, 8 и 9, или, наконец, полного объема книги.

Методической особенностью книги является то, что ряд Фурье вводится как представление сигнала в ортогональном пространстве. Это сделано из-за растущего значения геометрического представления сигналов в теории передачи информации. Однако в последующем изложении геометрическое представление не используется, поэтому первые три раздела 1.1—1.3 могут быть опущены.

Изучение материала книги не требует привлечения каких-либо дополнительных источников. Предполагается, что часть студентов не имеет знаний по теории вероятностей. Минимальные сведения из теории вероятностей, которые требуются в гл. 9 (посвященной цифровой связи), приводятся в этой же главе.

Пеория, Иллинойс
Январь 1968

Б. П. Латхи

ГЛАВА

1

Анализ сигналов

1.1. Введение

Существует много различных способов связи. Так, два человека могут связываться друг с другом, пользуясь речью, жестами или графическими символами. В далеком прошлом связь на большие расстояния осуществлялась с помощью звуковых сигналов барабана, дыма костров, почтовых голубей и световых лучей. Теперь связь на большие расстояния реализуется в основном с помощью электрических сигналов. Объясняется это тем, что электрические сигналы можно передавать на огромные расстояния (теоретически на любые расстояния в пределах Вселенной) с очень большой скоростью (около 3×10^8 *м/сек*). В этой книге рассматривается связь только с помощью электрических сигналов.

Для инженера интерес представляют в основном способы эффективной связи, при которых сообщения можно передавать с наибольшей скоростью и наименьшими ошибками. Количественный анализ этих вопросов приводится в книге. Полезен, однако, качественный анализ факторов, ограничивающих скорость передачи сигналов. Рассмотрим передачу символов (например, букв английского алфавита) с помощью электрических сигналов определенной формы.

В процессе передачи сигналы подвергаются воздействию помех, порождаемых многочисленными природными явлениями и возникающих в различных устройствах. По этому признаку помехи можно разделить на естественные и искусственные. Источниками искусственных помех являются, например, контактные переключатели, системы зажигания, устройства флюоресцентного свечения, непрерывно излучающие случайные шумовые сигналы. Естественными источниками помех служат такие природные явления, как молнии, электрические бури, солнечные и внутригалактические излучения. Весьма существенным во всякого рода электрических системах является флуктуационный шум (тепловой в резисторах и дробовой в активных электронных приборах). При передаче по каналу связи несущие информацию сигналы искажаются и вследствие этого могут быть неверно приняты. Для устранения указанного явления приходится увеличивать мощности передаваемых сиг-

налов, чтобы поддерживать необходимое отношение мощности сигнала к мощности шума. Отношение сигнал/шум S/N является важным параметром системы связи.

Рассмотрим теперь увеличение скорости передачи сообщений посредством сжатия сигналов во времени для передачи большего числа сообщений за данный отрезок времени. При уменьшении длительности сигналов их мгновенные значения изменяются быстрее, т. е. увеличиваются их частоты. Следовательно, при сжатии сигналов возникает проблема передачи сигналов с более высокими частотами, что требует расширения полосы пропускания канала связи. Таким образом, скорость передачи информации можно увеличить за счет расширения полосы пропускания канала. Поэтому чем быстрее и точнее должно быть передано сообщение, тем больше должны быть отношение мощности сигнала к мощности шума и полоса пропускания канала.

Полученные выводы являются несколько неожиданными. Ниже будет показано, что для достижения заданной скорости передачи при требуемой точности полоса пропускания и отношение сигнал/шум могут обмениваться. Эта закономерность выражается формулой Шеннона-Хартли:

$$C = B \log \left(1 + \frac{S}{N}\right),$$

где C — емкость канала или скорость передачи сообщений (рассматривается ниже), B — ширина полосы пропускания канала, *гц*. При заданной скорости C можно увеличить B и уменьшить S/N и наоборот.

Для изучения систем связи необходимо познакомиться с различными способами представления сигналов. Эта глава как раз и посвящена анализу сигналов.

1.2. Аналогия между векторами и сигналами

ВЕКТОРЫ

Любая задача лучше воспринимается, если ее можно связать с каким-либо известным явлением. Аналогии особенно полезны при изучении абстрактных задач, если последние можно сопоставить какому-либо конкретному явлению. В этом случае новую задачу легче понять, используя знания о соответствующем ей явлении. Поэтому прежде чем перейти к анализу сигналов, рассмотрим аналогию между векторами и сигналами. Остановимся кратко на свойствах векторов.

Вектор характеризуется величиной и направлением. Рассмотрим два вектора $\mathbf{V}_1$ и $\mathbf{V}_2$ (рис. 1.1). Пусть составляющая вектора $\mathbf{V}_1$, совпадающая с направлением вектора $\mathbf{V}_2$, есть $C_{12}\mathbf{V}_2$. Каков физический смысл составляющей одного вектора, совпадающей с направлением другого? Геометрически составляющая вектора $\mathbf{V}_1$,

совпадающая с направлением вектора $\mathbf{V}_2$, находится, если из конца вектора $\mathbf{V}_1$ опустить перпендикуляр на $\mathbf{V}_2$, как это показано на рис. 1.1. Вектор $\mathbf{V}_1$ в этом случае можно выразить через вектор $\mathbf{V}_2$ следующим образом:

$$\mathbf{V}_1 = C_{12}\mathbf{V}_2 + \mathbf{V}_e. \tag{1.1а}$$

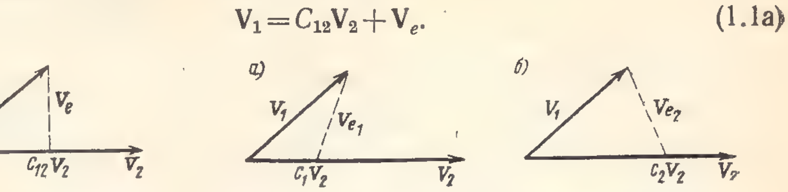


Рис. 1.1. Проекция одного вектора на направление другого

Рис. 1.2. Варианты аппроксимации векторов

Однако это не единственный способ выражения вектора $\mathbf{V}_1$ через $\mathbf{V}_2$. На рис. 1.2 показаны два из бесконечного множества возможных вариантов такого выражения. Так, согласно рис. 1.2*а*

$$\mathbf{V}_1 = C_1\mathbf{V}_2 + \mathbf{V}_{e1}, \tag{1.1б}$$

а согласно рис. 1.2*б*

$$\mathbf{V}_1 = C_2\mathbf{V}_2 + \mathbf{V}_{e2}. \tag{1.1в}$$

В каждом из этих выражений $\mathbf{V}_1$ представляется с помощью $\mathbf{V}_2$ и дополнительного вектора, который можно назвать вектором ошибки. Если требуется аппроксимировать вектор $\mathbf{V}_1$ вектором, совпадающим с направлением $\mathbf{V}_2$, то $\mathbf{V}_e$ есть вектор ошибки при этой аппроксимации. Например, (см. рис. 1.1), если $\mathbf{V}_1$ аппроксимируется величиной $C_{12}\mathbf{V}_2$, вектором ошибки такой аппроксимации является $\mathbf{V}_e$. Если $\mathbf{V}_1$ аппроксимируется вектором $C_1\mathbf{V}_2$ (см. рис. 1.2*а*), то вектор ошибки — $\mathbf{V}_{e1}$ и т. д. Что же является характерным для представления векторов на рис. 1.1? Из геометрического построения следует, что вектор ошибки на рис. 1.1 оказывается наименьшим. Теперь можно сформулировать количественное определение составляющей вектора, совпадающей с направлением другого вектора. Составляющая вектора $\mathbf{V}_1$, совпадающая с направлением вектора $\mathbf{V}_2$, представляется величиной $C_{12}\mathbf{V}_2$, где C_{12} выбирается так, чтобы вектор ошибки был минимальным.

Объясним теперь физический смысл составляющей одного вектора, совпадающей с направлением другого. Очевидно, чем больше составляющая одного вектора в направлении другого, тем меньше различие в направлениях обоих векторов и тем меньше вектор ошибки. Если составляющая вектора $\mathbf{V}_1$ в направлении $\mathbf{V}_2$ есть $C_{12}\mathbf{V}_2$, то значение C_{12} показывает степень подобия двух векторов. Если C_{12} равно нулю, данный вектор не имеет составляющей на направление другого вектора и, следовательно, оба вектора взаимно перпендикулярны. Эти векторы называют *ортогональными*. Та-

ким образом, ортогональные векторы независимы. Если векторы ортогональны, параметр C_{12} равен нулю.

Для удобства определим скалярное произведение двух векторов **A** и **B** как $\mathbf{A}\cdot\mathbf{B}=AB\cos\theta$, где θ — угол между векторами **A** и **B**. Из определения следует, что $\mathbf{A}\cdot\mathbf{B}=\mathbf{B}\cdot\mathbf{A}$. В соответствии с этим обозначением:

составляющая **A**,
имеющая направление $\mathbf{B}=A\cos\theta=\mathbf{A}\cdot\mathbf{B}/B$,
составляющая **B**,
имеющая направление $\mathbf{A}=B\cos\theta=\mathbf{A}\cdot\mathbf{B}/A$.

В частности,

составляющая $\mathbf{V}_1$,
имеющая направление $\mathbf{V}_2=\mathbf{V}_1\cdot\mathbf{V}_2/V_2=C_{12}\mathbf{V}_2$.

Следовательно,

$$C_{12}=\mathbf{V}_1\cdot\mathbf{V}_2/V_2^2=\mathbf{V}_1\cdot\mathbf{V}_2/\mathbf{V}_2\cdot\mathbf{V}_2. \quad (1.2)$$

Заметим, что если $\mathbf{V}_1$ и $\mathbf{V}_2$ ортогональны, то

$$\mathbf{V}_1\cdot\mathbf{V}_2=0, \quad C_{12}=0. \quad (1.3)$$

СИГНАЛЫ

Понятия сравнения векторов и их ортогональности можно применить и к сигналам[1]. Рассмотрим два сигнала $f_1(t)$ и $f_2(t)$. Предположим, что необходимо аппроксимировать $f_1(t)$ с помощью $f_2(t)$ в определенном интервале $(t_1<t<t_2)$, т. е.

$$f_1(t)\approx C_{12}f_2(t) \quad (t_1<t<t_2). \quad (1.4)$$

Как выбрать C_{12}, чтобы аппроксимация была наилучшей? Очевидно, C_{12} нужно выбрать так, чтобы ошибка между действительной функцией и аппроксимированной на данном интервале $(t_1<<t<t_2)$ была минимальной. Определим *функцию ошибки* как

$$f_e(t)=f_1(t)-C_{12}f_2(t). \quad (1.5)$$

Один из возможных критериев минимизации ошибки $f_e(t)$ на интервале $(t_1<t<t_2)$ состоит в минимизации средней величины $f_e(t)$ на этом интервале: $\frac{1}{t_2-t_1}\int_{t_1}^{t_2}[f_1(t)-C_{12}f_2(t)]dt$. Этот критерий, однако, допускает большие положительные и отрицательные ошибки, которые могут компенсироваться в процессе усреднения и давать ложный результат, что ошибка равна нулю. Например, при аппроксимации функции $\sin t$ функцией $f(t)=0$ на

---

[1]) В дальнейшем термины «сигналы» и «функции» будут часто использоваться в одном и том же смысле. Сигнал есть функция времени. Однако существует различие между сигналами и функциями. Функция $f(t)$ может быть многозначной относительно переменной t. Физически существующий сигнал всегда является однозначной функцией t. Следовательно, под термином «функция» будем понимать однозначную функцию независимой переменной.

интервале $(0,2\pi)$ средняя ошибка равна нулю. Это приводит к неверному заключению, что $\sin t$ можно аппроксимировать нулем на интервале от 0 до 2π без какой-либо ошибки. Указанный недостаток можно устранить, если минимизировать не саму ошибку, а среднее значение ее квадрата. Пусть среднее значение $f_e^2(t)$, называемое средней квадратической ошибкой, обозначено через

$$\varepsilon = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f_e^2(t) dt = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [f_1(t) - C_{12} f_2(t)]^2 dt. \qquad (1.6)$$

Для того чтобы найти величину C_{12}, которая минимизирует ε, необходимо положить

$$d\varepsilon / dC_{12} = 0, \qquad (1.7)$$

т. е.

$$\frac{d}{dC_{12}} \left\{ \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} [f_1(t) - C_{12} f_2(t)]^2 \, dt \right\} = 0. \qquad (1.8)$$

Изменив порядок интегрирования и дифференцирования, получим

$$\frac{1}{t_2 - t_1} \left[\int_{t_1}^{t_2} \frac{d}{dC_{12}} f_1^2(t) \, dt - 2 \int_{t_1}^{t_2} f_1(t) f_2(t) \, dt + 2C_{12} \int_{t_1}^{t_2} f_2^2(t) dt \right] = 0. \qquad (1.9)$$

Первый интеграл равен нулю, и, следовательно, по формуле (1.9)

$$C_{12} = \int_{t_1}^{t_2} f_1(t) f_2(t) dt \Big/ \int_{t_1}^{t_2} f_2^2(t) \, dt. \qquad (1.10)$$

Отметим сходство между ф-лами (1.10) и (1.2). По аналогии с векторами сигнал $f_1(t)$ имеет составляющую сигнала $f_2(t)$, и величина этой составляющей равна C_{12}. Если значение C_{12} равно нулю, то сигнал $f_1(t)$ не имеет составляющей сигнала $f_2(t)$, т. е. функции ортогональны на интервале (t_1, t_2). Следовательно, две функции $f_1(t)$ и $f_2(t)$ *ортогональны* на интервале (t_1, t_2), если

$$\int_{t_1}^{t_2} f_1(t) f_2(t) \, dt = 0. \qquad (1.11)$$

Обратим внимание на сходство ф-лы (1.11), полученной для ортогональных функций, с ф-лой (1.3) для ортогональных векторов.

Нетрудно показать, что функции $\sin m\omega_0 t$ и $\sin n\omega_0 t$ ортогональны на любом интервале $\left(t_0, t_0 + \frac{2\pi}{\omega_0}\right)$ при произвольных значениях m и n. Рассмотрим интеграл

$$I=\int\limits_{t_0}^{t_0+\frac{2\pi}{\omega_0}} \sin n\,\omega_0 t \sin m\,\omega_0 t=\int\limits_{t_0}^{t_0+\frac{2\pi}{\omega_0}} \frac{1}{2}\left[\cos(n-m)\omega_0 t-\cos(n+m)\omega_0 t\right]\,dt=$$

$$=\frac{1}{2\omega_0}\left[\frac{1}{n-m}\sin(n-m)\,\omega_0 t-\frac{1}{n+m}\sin(n+m)\,\omega_0 t\right]_{t_0}^{t_0+\frac{2\pi}{\omega_0}}.$$

Так как n и m — целые числа, то $(n-m)$ и $(n+m)$ — также целые числа. В этом случае интеграл I равен нулю, что свидетельствует об ортогональности функций. Подобным образом можно показать, что ортогональными являются функции $\sin n\,\omega_0 t$ и $\cos m\,\omega_0 t$, а также $\cos n\,\omega_0 t$ и $\cos m\,\omega_0 t$.

Пример 1.1.

Прямоугольная функция $f(t)$ (рис. 1.3) может быть записана в виде

$$f(t)=\begin{cases}1 \text{ при } (0<t<\pi),\\ -1 \text{ при } (\pi<t<2\pi).\end{cases}$$

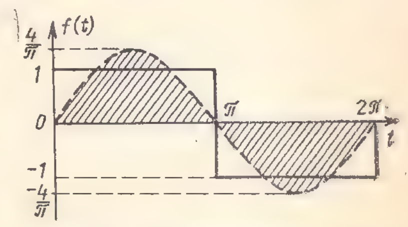


Рис. 1.3. Аппроксимация прямоугольной функции синусоидальной

Необходимо аппроксимировать эту функцию на интервале $(0, 2\pi)$ синусоидой $\sin t$ так, чтобы средняя квадратическая ошибка была минимальной.

Будем аппроксимировать функцию $f(t)$ на интервале $(0, 2\pi)$ следующим образом: $f(t)\approx C_{12}\sin t$.

Найдем оптимальное значение C_{12}, при котором средняя квадратическая ошибка аппроксимации минимальна. В соответствии с ф-лой (1.10) минимум средней квадратической ошибки получается, если

$$C_{12}=\frac{\int\limits_0^{2\pi} f(t)\sin t\,dt}{\int\limits_0^{2\pi}\sin^2 t\,dt}=\frac{1}{\pi}\left[\int\limits_0^{\pi}\sin t\,dt+\int\limits_{\pi}^{2\pi}-\sin t\,dt\right]=\frac{4}{\pi}.$$

Таким образом, выражение $f(t)\approx\frac{4}{\pi}\sin t$ представляет наилучшую аппроксимацию сигнала $f(t)$ функцией $\sin t$. По аналогии с векторами можно сказать, что прямоугольный сигнал $f(t)$, изображенный на рис. 1.3, содержит синусоидальную функцию с амплитудой $4/\pi$.

Что же означает ортогональность двух функций? В случае векторов ортогональность означает, что один вектор не содержит компоненты, направленной вдоль другого вектора. Аналогично функция не содержит в своем составе компоненты, имеющей форму ортогональной ей функции. Если аппроксимировать функцию ортогональной ей функцией, ошибка будет больше самой аппроксимируемой функции. Поэтому лучше аппроксимировать функцию не ортогональной, а нулевой функцией $f(t)=0$. Следовательно, в этом случае оптимальное значение $C_{12}=0$.

ГРАФИЧЕСКАЯ ОЦЕНКА СОСТАВЛЯЮЩЕЙ ОДНОЙ ФУНКЦИИ В ДРУГОЙ

На основании ф-лы (1.10) можно графически вычислить составляющую одной функции в другой. Допустим, известно графическое представление функций $f_1(t)$ и $f_2(t)$ и требуется оценить составляющую функции $f_2(t)$, содержащуюся в сигнале $f_1(t)$ на интервале $(0, T)$. Известно, что искомая составляющая равна $C_{12}f_2(t)$, где $C_{12} = \int\limits_0^T f_1(t)\, f_2(t)dt \Big/ \int\limits_0^T f_2^2(t)\, dt$. Интеграл в числителе этой формулы можно найти, перемножив две функции и определив площадь под результирующей кривой, как это показано на рис. 1.4. Интеграл в знаменателе можно вычислить, если аналогично найти площадь под функцией $[f_2(t)]^2$.

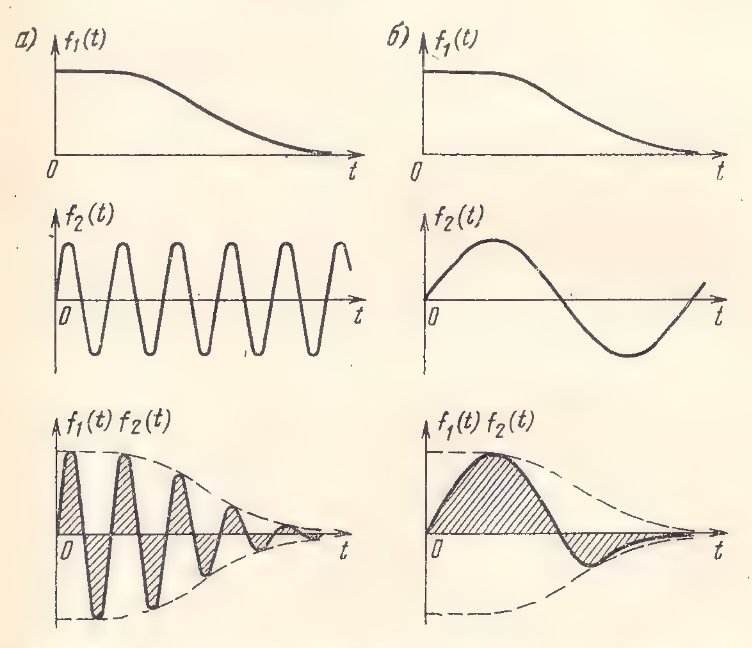


Рис. 1.4. Графическое определение составляющей сигнала $f_2(t)$ в $f_1(t)$.

Очевидно, если $f_1(t)$ изменяется значительно медленнее функции $f_2(t)$, площадь под кривой $f_1(t)f_2(t)$ будет небольшой, так как положительные и отрицательные участки ее примерно равны и компенсируются (рис. 1.4*а*). Это означает, что функция $f_1(t)$ содержит весьма малую составляющую $f_2(t)$. Если же, однако, $f_1(t)$ изменяется примерно с той же скоростью, что и $f_2(t)$, площадь под кривой $f_1(t)f_2(t)$ достаточно велика (рис. 1.4*б*) и, следова-

тельно, $f_1(t)$ содержит значительную по величине составляющую функции $f_2(t)$. Этот результат очевиден, так как если две функции изменяются с примерно одинаковой скоростью, можно ожидать значительного сходства между ними и, следовательно, $f_1(t)$ будет содержать бо́льшую составляющую функции $f_2(t)$.

ОРТОГОНАЛЬНОЕ ВЕКТОРНОЕ ПРОСТРАНСТВО

Можно расширить аналогию между векторами и сигналами, рассмотрев трехмерное векторное пространство (рис. 1.5). Обозначим единичный вектор вдоль оси x через $\mathbf{a}_x$, вдоль оси y через $\mathbf{a}_y$ и вдоль оси z через $\mathbf{a}_z$. Так как длины векторов $\mathbf{a}_x$, $\mathbf{a}_y$ и $\mathbf{a}_z$ равны единице, для любого вектора $\mathbf{A}$ можно записать:

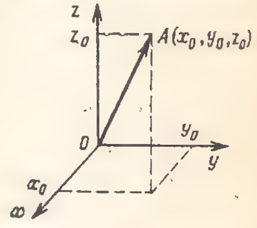


Рис. 1.5. Трехмерное векторное пространство

составляющая вектора $\mathbf{A}$ вдоль оси $x = \mathbf{A} \cdot \mathbf{a}_x$,
составляющая вектора $\mathbf{A}$ вдоль оси $y = \mathbf{A} \cdot \mathbf{a}_y$,
составляющая вектора $\mathbf{A}$ вдоль оси $z = \mathbf{A} \cdot \mathbf{a}_z$.

Вектор $\mathbf{A}$, проведенный из начала координат к произвольной точке пространства (x_0, y_0, z_0), имеет составляющие x_0, y_0 и z_0 вдоль осей x, y и z соответственно. Теперь вектор $\mathbf{A}$ можно выразить с помощью составляющих по трем взаимно перпендикулярным осям:

$$\mathbf{A} = x_0\mathbf{a}_x + y_0\mathbf{a}_y + z_0\mathbf{a}_z.$$

Так как $\mathbf{a}_x$, $\mathbf{a}_y$, $\mathbf{a}_z$ взаимно перпендикулярны,

$$\left.\begin{aligned} \mathbf{a}_x \cdot \mathbf{a}_y = \mathbf{a}_y \cdot \mathbf{a}_z = \mathbf{a}_z \cdot \mathbf{a}_x = 0 \\ \mathbf{a}_x \cdot \mathbf{a}_x = \mathbf{a}_y \cdot \mathbf{a}_y = \mathbf{a}_z \cdot \mathbf{a}_z = 1 \end{aligned}\right\}. \tag{1.12}$$

Это свойство можно выразить более лаконично:

$$\mathbf{a}_m \cdot \mathbf{a}_n = \begin{cases} 0 \text{ при } m \neq n, \\ 1 \text{ при } m = n, \end{cases} \tag{1.13}$$

где m, n — любые индексы x, y или z.

Следует отметить, что если система координат имеет только оси x, y, то она недостаточна для точного выражения произвольного вектора $\mathbf{A}$ через составляющие вдоль этих осей. Эта система позволяет выразить две составляющие вектора $\mathbf{A}$. Для того чтобы можно было выразить произвольный вектор $\mathbf{A}$ через его координатные составляющие, необходимо иметь полную систему координат. В данном случае должны быть три координатных оси.

Полученные выше выводы можно обобщить на случай n-мерного пространства. Физически такое пространство в природе не

существует. Тем не менее имеется множество задач, которые можно рассматривать как n-мерные. Например, линейное уравнение с n независимыми переменными можно рассматривать как вектор, выраженный через свои составляющие по n взаимно перпендикулярным осям координат. Если единичные векторы вдоль этих n взаимно перпендикулярных осей обозначить через $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, а составляющие произвольного вектора $\mathbf{A}$ в n-мерном пространстве соответственно через $C_1, C_2, \ldots, C_n$, то

$$\mathbf{A} = C_1\mathbf{x}_1 + C_2\mathbf{x}_2 + C_3\mathbf{x}_3 + \ldots + C_n\mathbf{x}_n. \tag{1.14}$$

Все векторы $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ взаимно ортогональны, и для представления любого произвольного вектора $\mathbf{A}$ согласно ф-ле (1.14) ряд должен быть полным. Условие ортогональности означает, что скалярное произведение любых двух векторов $\mathbf{x}_n$ и $\mathbf{x}_m$ должно быть равно нулю, а скалярное произведение любого вектора самого на себя должно быть равно единице. Условие (1.13) можно записать как

$$\mathbf{x}_m \cdot \mathbf{x}_n = \begin{cases} 0 \text{ при } m \neq n, \\ 1 \text{ при } m = n. \end{cases} \tag{1.15}$$

Постоянные $C_1, C_2, C_3, \ldots, C_n$ в ф-ле (1.14) представляют собой значения составляющих вектора $\mathbf{A}$ вдоль соответствующих векторов $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n$, следовательно,

$$C_r = \mathbf{A} \cdot \mathbf{x}_r. \tag{1.16}$$

Этот же результат можно получить, определив скалярное произведение обеих частей равенства (1.14) с вектором $\mathbf{x}_r$:

$$\mathbf{A} \cdot \mathbf{x}_r = C_1\mathbf{x}_1 \cdot \mathbf{x}_r + C_2\mathbf{x}_2 \cdot \mathbf{x}_r + \ldots + C_r\mathbf{x}_r \cdot \mathbf{x}_r + \ldots + C_n\mathbf{x}_n \cdot \mathbf{x}_r. \tag{1.17}$$

Согласно ф-ле (1.15) все слагаемые вида $C_j\mathbf{x}_j \cdot \mathbf{x}_r$ $(j \neq r)$ в правой части ф-лы (1.17) равны нулю. Поэтому

$$\mathbf{A} \cdot \mathbf{x}_r = C_r\mathbf{x}_r \cdot \mathbf{x}_r = C_r. \tag{1.18}$$

Назовем ряд векторов $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ *ортогональным векторным пространством.* В общем случае скалярное произведение $\mathbf{x}_m \cdot \mathbf{x}_n$ может быть равно не единице, а некоторой постоянной k_m. Если $k_m = 1$, ряд называется нормализованным ортогональным рядом, или *ортонормальным векторным пространством.* Поэтому в общем случае для ортогонального векторного пространства $\{\mathbf{x}_r\} \ldots (r = 1, 2, \ldots, n)$ имеем

$$\mathbf{x}_m \cdot \mathbf{x}_n = \begin{cases} 0 & \text{при } m \neq n, \\ k_m & \text{при } m = n. \end{cases} \tag{1.19}$$

Формулу (1.18) для ортогонального векторного пространства можно записать в виде

$$\mathbf{A} \cdot \mathbf{x}_r = C_r\mathbf{x}_r \cdot \mathbf{x}_r = C_r k_r \quad (C_r = \mathbf{A} \cdot \mathbf{x}_r / k_r). \tag{1.20}$$

Теперь кратко сформулируем результаты. Если ортогональное векторное пространство $\{\mathbf{x}_r\} \ldots (r=1, 2\ldots)$ является полным, любой вектор $\mathbf{F}$ можно представить как

$$\mathbf{F}=C_1\mathbf{x}_1+C_2\mathbf{x}_2+\ \ldots+C_r\mathbf{x}_r+\ \ldots, \tag{1.21}$$

где

$$C_r=\frac{\mathbf{F}\cdot\mathbf{x}_r}{k_r}=\frac{\mathbf{F}\cdot\mathbf{X}_r}{\mathbf{x}_r\cdot\mathbf{x}_r}. \tag{1.22}$$

ОРТОГОНАЛЬНОЕ СИГНАЛЬНОЕ ПРОСТРАНСТВО

Применим к анализу сигналов некоторые представления векторного пространства. Выше было показано, что любой вектор можно выразить в виде суммы его составляющих по взаимно перпендикулярным векторам, если эти векторы образуют полную систему координат. Поэтому следует ожидать, что и любую функцию $f(t)$ можно выразить суммой взаимно ортогональных функций, если они образуют полный ряд. Покажем, что это действительно возможно.

Аппроксимация функции рядом взаимно ортогональных функций. Рассмотрим ряд из n функций $g_1(t), g_2(t), \ldots, g_n(t)$, которые ортогональны друг другу на интервале (t_1, t_2), т. е.

$$\int\limits_{t_1}^{t_2} g_i(t)\, g_k(t)\, dt=0 \quad (i \neq k) \tag{1.23а}$$

и пусть

$$\int\limits_{t_1}^{t_2} g_i^2(t)\, dt=K_i. \tag{1.23б}$$

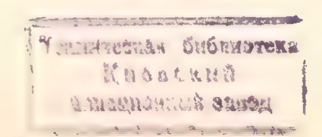
Будем аппроксимировать произвольную функцию $f(t)$ на интервале (t_1, t_2) линейной комбинацией этих n взаимно ортогональных функций:

$$f(t) \approx C_1 g_1(t)+C_2 g_2(t)+\ \ldots+C_k g_k(t)+\ \ldots+C_n g_n(t)=\sum_{r=1}^{n} C_r g_r(t). \tag{1.24}$$

Чтобы получить наилучшую аппроксимацию, необходимо найти значения постоянных $C_1, C_2, \ldots, C_n$, при которых средняя квадратическая величина функции $f_e(t)$ минимальна. По определению

$f_e(t)=f(t)-\sum\limits_{r=1}^{n} C_r g_r(t)$ и

$$\varepsilon=\frac{1}{t_2-t_1}\int\limits_{t_1}^{t_2}\left[f(t)-\sum_{r=1}^{n} C_r g_r(t)\right]^2 dt. \tag{1.25}$$



Из ф-лы (1.25) следует, что ε является функцией $C_1, C_2, \ldots C_n$, и для ее минимизации необходимо положить $\partial\varepsilon/\partial C_1 = \partial\varepsilon/\partial C_2 = \ldots = \partial\varepsilon/\partial C_i = \ldots = \partial\varepsilon/\partial C_n = 0$. Рассмотрим уравнение

$$\partial\varepsilon/\partial C_i = 0. \qquad (1.26)$$

Так как $(t_2 - t_1)$ — величина постоянная, ф-лу (1.25) можно переписать в виде

$$\frac{\partial}{\partial C_i}\left\{\int_{t_1}^{t_2}\left[f(t) - \sum_{r=1}^{n} C_r g_r(t)\right]^2 dt\right\} = 0. \qquad (1.27)$$

Если возвести в квадрат выражение в квадратных скобках под знаком интеграла, то в силу ортогональности все слагаемые вида $\int g_i(t) g_k(t)\, dt$ будут равны нулю. Таким образом, производная всех слагаемых, не содержащих C_i, равна нулю, т. е.

$$\frac{\partial}{\partial C_i}\int_{t_1}^{t_2} f^2(t)\, dt = \frac{\partial}{\partial C_i}\int_{t_1}^{t_2} C_r^2 g_r^2(t)\, dt = \frac{\partial}{\partial C_i}\int_{t_1}^{t_2} C_r f(t) g_r(t) dt = 0.$$

В результате в ф-ле (1.27) остается только два не равных нулю слагаемых:

$$\frac{\partial}{\partial C_i}\int_{t_1}^{t_2}\left[-2C_i f(t) g_i(t) + C_i^2 g_i^2(t)\right] dt = 0. \qquad (1.28)$$

Изменяя в ф-ле (1.28) порядок интегрирования и дифференцирования, получаем:

$$2\int_{t_1}^{t_2} f(t) g_i(t)\, dt = 2C_i \int_{t_1}^{t_2} g_i^2(t) dt.$$

Следовательно,

$$C_i = \frac{\int_{t_1}^{t_2} f(t) g_i(t)\, dt}{\int_{t_1}^{t_2} g_i^2(t)\, dt} = \frac{1}{K_i}\int_{t_1}^{t_2} f(t) g_i(t)\, dt. \qquad (1.29)$$

Итак, показано, что произвольную функцию $f(t)$ можно аппроксимировать на интервале (t_1, t_2) рядом из n взаимно ортогональных на этом интервале функций $g_1(t)$, $g_2(t)$... с помощью их линейной комбинации (1.24).

Для достижения наилучшей в смысле минимума средней квадратической ошибки аппроксимации на заданном интервале коэффициенты $C_1, C_2, \ldots, C_n$ должны выбираться в соответствии с ф-лой (1.29).

Вычисление средней квадратической ошибки. Найдем теперь величину ε, если коэффициенты $C_1, C_2, \ldots, C_n$ выбраны в соответствии с ф-лой (1.29). По определению

$$\varepsilon = \frac{1}{t_2 - t_1}\int\limits_{t_1}^{t_2}\left[f(t) - \sum_{r=1}^{n} C_r g_r(t)\right]^2 dt =$$

$$= \frac{1}{t_2 - t_1}\left[\int\limits_{t_1}^{t_2} f^2(t)\,dt + \sum_{r=1}^{n} C_r^2 \int\limits_{t_1}^{t_2} g_r^2(t)dt - 2\sum_{r=1}^{n} C_r \int\limits_{t_1}^{t_2} f(t)\,g_r(t)\,dt\right]. \quad (1.30)$$

Из ф-лы (1.29) следует, что

$$\int\limits_{t_1}^{t_2} f(t)\,g_r(t)\,dt = C_r \int\limits_{t_1}^{t_2} g_r^2(t)\,dt = C_r K_r. \quad (1.31)$$

Подставив ф-лу (1.31) в (1.30), получим

$$\varepsilon = \frac{1}{t_2 - t_1}\left[\int\limits_{t_1}^{t_2} f^2(t)dt + \sum_{r=1}^{n} C_r^2 K_r - 2\sum_{r=1}^{n} C_r^2 K_r\right] =$$

$$= \frac{1}{t_2 - t_1}\left[\int\limits_{t_1}^{t_2} f^2(t)dt - \sum_{r=1}^{n} C_r^2 K_r\right] = \quad (1.32)$$

$$= \frac{1}{t_2 - t_1}\left[\int\limits_{t_1}^{t_2} f^2(t)\,dt - (C_1^2 K_1 + C_2^2 K_2 + \ \ldots + C_n^2 K_n\right]. \quad (1.33)$$

Выражение (1.33) позволяет определить величину средней квадратической ошибки.

Представление функции замкнутым или полным рядом взаимно ортогональных функций. Из ф-лы (1.33) следует, что при увеличении n, т. е. при аппроксимации $f(t)$ бо́льшим числом ортогональных функций, ошибка уменьшается. Но по определению ε является положительной величиной; следовательно, в пределе при бесконечном числе слагаемых сумма $\sum\limits_{r=1}^{\infty} C_r^2\ K_r$ сходится к интегралу $\int\limits_{t_1}^{t_2} f^2(t)dt$ и ошибка ε стремится к нулю. Таким образом,

$$\int\limits_{t_1}^{t_2} f^2(t)\,dt = \sum_{r=1}^{\infty} C_r^2 K_r. \quad (1.34)$$

В этом случае $f(t)$ представляется бесконечным рядом $f(t) = C_1 g_1(t) + C_2 g_2(t) + \ldots + C_r g_r(t) + \ldots$. Бесконечный ряд в правой части ф-лы (1.34) сходится к функции $f(t)$ так, что средняя квадратическая ошибка равна нулю. Такие ряды называются *сходящимися в среднем.* Заметим, что теперь представление $f(t)$ оказывается точным.

Множество функций $g_1(t)$, $g_2(t), \ldots, g_r(t)$, взаимно ортогональных на интервале (t_1, t_2), называют полным или замкнутым, если не существует такой функции $x(t)$, для которой справедливо выражение

$$\int_{t_1}^{t_2} x(t)\, g_k(t)\, dt = 0 \quad (k = 1, 2 \ldots).$$

Если же можно найти функцию $x(t)$, для которой этот интеграл равен нулю, то $x(t)$ является ортогональной каждой функции множества $\{g_r(t)\}$ и, следовательно, сама функция $x(t)$ является членом указанного множества. Очевидно, множество без $x(t)$ не будет полным. Таким образом, для множества функций $\{g_r(t)\}$ $(r=1, 2\ldots)$, взаимно ортогональных на интервале (t_1, t_2), справедливо соотношение:

$$\int_{t_1}^{t_2} g_m(t)\, g_n(t)\, dt = \begin{cases} 0 & \text{при } m \neq n, \\ K_m & \text{при } m = n. \end{cases} \tag{1.35}$$

Если это множество полное, то любую функцию $f(t)$ можно выразить рядом

$$f(t) = C_1 g_1(t) + C_2 g_2(t) + \ldots + C_r g_r(t) + \ldots, \tag{1.36}$$

где

$$C_r = \frac{\int_{t_1}^{t_2} f(t)\, g_r(t)\, dt}{K_r} = \frac{\int_{t_1}^{t_2} f(t)\, g_r(t)\, dt}{\int_{t_1}^{t_2} g_r^2(t)\, dt}. \tag{1.37}$$

Аналогия между векторами и сигналами вытекает из сравнения равенств (1.35) и (1.37) с (1.19) и (1.22). Любой вектор можно выразить суммой взаимно ортогональных векторов, если эти векторы образуют полное множество. Аналогично, любую функцию $f(t)$ можно выразить суммой взаимно ортогональных функций, если они составляют полное или замкнутое множество.

При сравнении векторов и сигналов скалярное произведение двух векторов соответствует интегралу от произведения двух сигналов, т. е.

$$\mathbf{A} \cdot \mathbf{B} \sim \int_{t_1}^{t_2} f_A(t)\, f_B(t)\, dt.$$

Это означает, что квадрату длины A вектора $\mathbf{A}$ соответствует интеграл от квадрата функции, т. е.

$$\mathbf{A} \cdot \mathbf{A} = A^2 \sim \int_{t_1}^{t_2} f_A^2(t) dt.$$

Если вектор выражается через взаимно ортогональные составляющие, квадрат его длины определяется суммой квадратов длин

составляющих векторов. Аналогичный результат справедлив и для сигналов, что иллюстрируется ф-лой (1.34) (теорема Парсеваля). Так как составляющие функции не ортонормальны, правая часть формулы имеет вид $\Sigma C_r^2 K_r$ вместо ΣC_r^2. Для ортонормального множества $K_r=1$. Равенство (1.34) соответствует случаю, когда вектор выражается через составляющие, направленные вдоль взаимно ортогональных векторов, квадраты длин которых равны $K_1, K_2, \ldots, K_r \ldots$

Разложение (1.36) показывает, что $f(t)$ состоит из функций $g_r(t)$ с коэффициентами C_r. Представление функции $f(t)$ бесконечным множеством ортогональных функций называют *обобщенным рядом Фурье функции* $f(t)$.

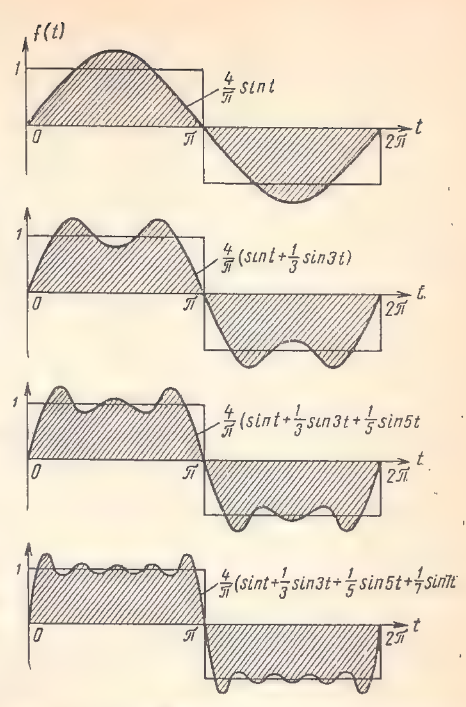


Рис. 1.6. Аппроксимация прямоугольной функции ортогональными

Пример 1.2

Рассмотрим, как и в примере 1.1 (см. рис. 1.3), прямоугольную функцию. Этот сигнал аппроксимировался функцией $\sin t$. Попытаемся улучшить аппроксимацию, воспользовавшись большим числом ортогональных функций: $f(t) \approx \approx C_1 \sin t + C_2 \sin 2t + \ldots + C_n \sin nt$.

Постоянные C_r вычислим по ф-ле (1.29):

$$C_r = \frac{\int\limits_0^{2\pi} f(t)\sin rt\,dt}{\int\limits_0^{2\pi} \sin^2 rt\,dt} = \frac{1}{\pi}\left(\int\limits_0^{\pi} \sin rt\,dt - \int\limits_{\pi}^{2\pi} \sin rt\,dt\right) =$$

$$\begin{cases} = \dfrac{4}{\pi r} & \text{для } r \text{ нечетного,} \\ = 0 & \text{для } r \text{ четного.} \end{cases}$$

В результате получим следующую аппроксимацию:

$$f(t) = \frac{4}{\pi}\left(\sin t + \frac{1}{3}\sin 3t + \frac{1}{5}\sin 5t + \frac{1}{7}\sin 7t + \ldots\right). \quad (1.38)$$

На рис. 1.6 показана прямоугольная функция до и после аппроксимации одним, двумя, тремя и четырьмя членами ряда (1.38). При заданном числе чле-

нов вида $\sin rt$ эти аппроксимации оптимальны, так как минимизируют среднюю квадратическую ошибку. С увеличением числа членов ряда аппроксимация улучшается, а средняя квадратическая ошибка уменьшается. При бесконечном числе членов ряда средняя квадратическая ошибка равна нулю [1]).

Оценим ошибку аппроксимации ε, воспользовавшись для этого ф-лой (1.33). В данном случае $t_2-t_1=2\pi$ и

$$f(t)=\begin{cases}1 \text{ при } (0<t<\pi),\\ -1 \text{ при } (\pi<t<2\pi).\end{cases}$$

Поэтому $\int\limits_0^{2\pi} f^2(t)\,dt=2\pi$,

$$C_r=\begin{cases}\dfrac{4}{\pi r} & \text{для } r \text{ нечетного}\\ 0 & \text{для } r \text{ четного,}\end{cases} \qquad K_r=\int\limits_0^{2\pi}\sin^2 rt\,dt=\pi.$$

При аппроксимации одним членом ряда

$$\varepsilon_1=\frac{1}{2\pi}\left[2\pi-\left(\frac{4}{\pi}\right)^2\pi\right]=0{,}19,$$

двумя членами ряда

$$\varepsilon_2=\frac{1}{2\pi}\left[2\pi-\left(\frac{4}{\pi}\right)^2\pi-\left(\frac{4}{3\pi}\right)^2\pi\right]=0{,}1,$$

тремя членами ряда

$$\varepsilon_3=\frac{1}{2\pi}\left[2\pi-\left(\frac{4}{\pi}\right)^2\pi-\left(\frac{4}{3\pi}\right)^2\pi-\left(\frac{4}{5\pi}\right)^2\pi\right]=0{,}0675,$$

четырьмя членами ряда

$$\varepsilon_4=\frac{1}{2\pi}\left[2\pi-\left(\frac{4}{\pi}\right)^2\pi-\left(\frac{4}{3\pi}\right)^2\pi-\left(\frac{4}{5\pi}\right)^2\pi-\left(\frac{4}{7\pi}\right)^2\pi\right]=0{,}051.$$

и т. д.

Легко видеть, что средняя квадратическая ошибка быстро уменьшается с увеличением числа членов аппроксимации.

Ортогональность комплексных функций. До сих пор рассматривались только действительные функции действительных переменных. Для комплексных функций $f_1(t)$ и $f_2(t)$ действительной переменной t можно показать, что $f_1(t)$ аппроксимируется функцией $f_2(t)$ на интервале $(t_1,\ t_2)$ как $f_1(t)\approx C_{12}f_2(t)$. Оптимальное значение C_{12}, минимизирующее среднюю квадратическую ошибку [2]):

$$C_{12}=\int\limits_{t_1}^{t_2} f_1(t)\,f_2^*(t)\,dt \Bigg/ \int\limits_{t_1}^{t_2} f_2(t)\,f_2^*(t)dt, \tag{1.39}$$

---

[1]) Ряды Фурье не удовлетворяют условиям сходимости в точках разрыва функции, и поэтому, несмотря на увеличение числа членов, аппроксимированная функция имеет выбросы в точках разрыва. Этот факт известен под названием *явления Гиббса.*

[2]) См., например, С. Мэзон, Г. Циммерман. Электронные цепи, сигналы и системы. ИИЛ, 1963, стр. 209—210.

где $f_2^*(t)$ — функция, комплексно-сопряженная с $f_2(t)$.

Из ф-лы (1.39) следует, что две комплексные функции $f_1(t)$ и $f_2(t)$ ортогональны на интервале (t_1, t_2), если

$$\int_{t_1}^{t_2} f_1(t) f_2^*(t)\, dt = \int_{t_1}^{t_2} f_1^*(t) f_2(t)\, dt = 0. \qquad (1.40)$$

Для множества комплексных функций $\{g_r(t)\}$ $(r=1, 2\ldots)$, взаимно ортогональных на интервале (t_1, t_2), справедливо соотношение

$$\int_{t_1}^{t_2} g_m(t) g_n^*(t)\, dt = \begin{cases} 0 & \text{при } m \neq n, \\ K_m & \text{при } m = n. \end{cases} \qquad (1.41)$$

Если это множество функций полное, функцию $f(t)$ можно аппроксимировать рядом (1.36), в котором

$$C_r = \frac{1}{K_r}\int_{t_1}^{t_2} f(t) g_r^*(t) dt. \qquad (1.42)$$

Если функции множества действительные, то $g_r^*(t) = g_r(t)$ и все результаты, полученные для комплексных функций, сводятся к ф-лам (1.35)—(1.37) для действительных функций.

1.3. Примеры ортогональных функций

РЯД ФУРЬЕ ИЗ ПОЛИНОМОВ ЛЕЖАНДРА

Представление функции на определенном интервале линейной комбинацией взаимно ортогональных функций называют рядом Фурье. Существует, однако, большое число рядов ортогональных функций, и, следовательно, любым из этих рядов можно выразить данную функцию. В векторном пространстве это аналогично представлению данного вектора в различных системах координат. Примерами множества ортогональных функций могут служить тригонометрические и комплексные экспоненциальные функции, полиномы Лежандра и Якоби. Бесселевы функции также являются специальным типом ортогональных функций[1]).

Полиномы Лежандра $P_n(x)$ $(n=0, 1, 2\ldots)$ образуют полное множество функций, ортогональных на интервале $(-1<t<1)$. Вид полиномов определяется формулой Родрига:

$$P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n \quad (n=0,1,2\ \ldots).$$

---

[1]) Бесселевы функции ортогональны с соответствующей весовой функцией. См., например, W. Kaplan. Advanced Calculus, Addison-Wesley, Reading, Mass., 1953, и Н. Н. Лебедев. Специальные функции и их приложения. Физматгиз, 1963.

Из этой формулы следует, что

$$\left.\begin{array}{ll} P_0(t)=1, & P_1(t)=t \\ P_2(t)=\frac{3}{2}t^2-\frac{1}{2}, & P_3(t)=\frac{5}{2}t^3-\frac{3}{2}t \end{array}\right\} \qquad (1.43)$$

и т. д.

Ортогональность этих полиномов можно проверить, убедившись в выполнении соотношения

$$\int_{-1}^{1} P_m(t)\,P_n(t)\,dt = \begin{cases} 0 & \text{при } m \neq n, \\ \dfrac{2}{2m+1} & \text{при } m=n. \end{cases} \qquad (1.44)$$

Функцию $f(t)$ можно выразить на интервале $(-1<t<1)$ с помощью полиномов Лежандра следующим образом:

$$f(t)=C_0P_0(t)+C_1P_1(t)+\ \ldots$$

где

$$C_r=\frac{\int\limits_{-1}^{1} f(t)\,P_r(t)\,dt}{\int\limits_{-1}^{1} P_r^2(t)\,dt}=\frac{2r+1}{2}\int\limits_{-1}^{1} f(t)P_r(t)\,dt. \qquad (1.45)$$

Отметим, что хотя представление с помощью полиномов Лежандра справедливо для интервала от -1 до 1, его можно распространить на любую область путем замены переменных.

Пример 1.3

Рассмотрим прямоугольную функцию (рис. 1.7). Эту функцию можно представить рядом Фурье из полиномов Лежандра

$$f(t)=C_0\,P_0(t)+C_1\,P_1(t)+\ldots+C_r\,P_r(t)+\ldots$$

Коэффициенты C_0, C_1, C_2, ..., C_r и т. д. находятся по ф-ле (1.45):

Рис. 1.7. Прямоугольная функция

$$C_0=\frac{1}{2}\int\limits_{-1}^{1} f(t)\,dt=0,$$

$$C_1=\frac{3}{2}\int\limits_{-1}^{1} t\,f(t)\,dt=\frac{3}{2}\left(\int\limits_{-1}^{0} t\,dt-\int\limits_{-1}^{0} t\,dt\right)=-\frac{3}{2},$$

$$C_2=\frac{5}{2}\int\limits_{-1}^{1} f(t)\left(\frac{3}{2}t^2-\frac{1}{2}\right)dt=\frac{5}{2}\left[\int\limits_{-1}^{0}\left(\frac{3}{2}t^2-\frac{1}{2}\right)dt+\right.$$

$$\left.+\int\limits_{0}^{1}-\left(\frac{3}{2}t^2-\frac{1}{2}\right)dt\right]=0.$$

Можно показать, что и остальные коэффициенты с четными индексами равны нулю, т. е.

$$C_4 = C_6 = \ldots = 0,$$

$$C_3 = \frac{7}{2}\int\limits_{-1}^{1} f(t)\left(\frac{5}{2}t^3 - \frac{3}{2}t\right)dt =$$

$$= \frac{7}{2}\left[\int\limits_{-1}^{0}\left(\frac{5}{2}t^3 - \frac{3}{2}t\right)dt + \int\limits_{0}^{1} -\left(\frac{5}{2}t^3 - \frac{3}{2}t\right)dt\right] = \frac{7}{8}.$$

Аналогично вычисляются коэффициенты C_5, C_7 ... и т. д. Окончательно имеем

$$f(t) = -\frac{3}{2}t + \frac{7}{8}\left(\frac{5}{2}t^3 - \frac{3}{2}t\right) + \ldots$$

ТРИГОНОМЕТРИЧЕСКИЙ РЯД ФУРЬЕ

Ранее было показано, что функции $\sin\omega_0 t$, $\sin 2\omega_0 t$ и т. д. образуют множество ортогональных функций на интервале $\left(t_0, t_0 + \frac{2\pi}{\omega_0}\right)$. Однако указанное множество не является полным. Это следует хотя бы из того, что функция $\cos n\omega_0 t$ является ортогональной $\sin m\omega_0 t$ на том же самом интервале. Следовательно, чтобы множество было полным, оно должно наряду с синусоидальными включать и косинусоидальные функции.

Можно показать, что множество, включающее функции $\cos n\omega_0 t$ и $\sin n\omega_0 t$ $(n=0, 1, 2...)$, является полным и ортогональным. Таким образом, любую функцию $f(t)$ можно представить на интервале $\left(t_0, t_0 + \frac{2\pi}{\omega_0}\right)$ рядом

$$f(t) = a_0 + a_1\cos\omega_0 t + a_2\cos 2\omega_0 t + \ldots + a_n\cos n\omega_0 t + \ldots +$$
$$+ b_1\sin\omega_0 t + b_2\sin 2\omega_0 t + \ldots + b_n\sin n\omega_0 t + \ldots$$

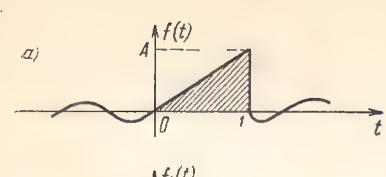
Введем обозначение $T = \frac{2\pi}{\omega_0}$ и перепишем это равенство в виде

$$f(t) = a_0 + \sum_{n=1}^{\infty}(a_n\cos n\omega_0 t + b_n\sin n\omega_0 t), \quad t_0 < t < t_0 + T \qquad (1.46)$$

Равенство (1.46) — тригонометрическая форма ряда Фурье. Коэффициенты a_n и b_n вычисляются по формулам:

$$a_n = \int\limits_{t_0}^{t_0+T} f(t)\cos n\omega_0 t\,dt \Bigg/ \int\limits_{t_0}^{t_0+T}\cos^2 n\omega_0 t\,dt, \qquad (1.47а)$$

$$b_n = \int\limits_{t_0}^{t_0+T} f(t)\sin n\omega_0 t\,dt \Bigg/ \int\limits_{t_0}^{t_0+T}\sin^2 n\omega_0 t\,dt. \qquad (1.47б)$$



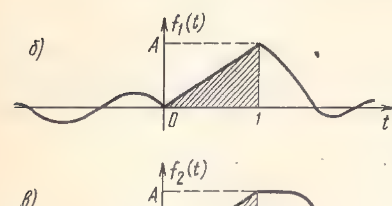


Рис. 1.8. Разложение треугольной функции в ряд Фурье

Если в ф-ле (1.47) положить $n=0$, то

$$a_0=\frac{1}{T}\int\limits_{t_0}^{t_0+T} f(t)\,dt. \qquad (1.48а)$$

Нетрудно найти, что $\int\limits_{t_0}^{t_0+T}\cos^2 \times$ $\times\, n\,\omega_0\, t dt=\int\limits_{t_0}^{t_0+T}\sin^2 n\,\omega_0\, t dt=T/2.$

Следовательно,

$$a_n=\frac{2}{T}\int\limits_{t_0}^{t_0+T} f(t)\cos n\,\omega_0\, t dt, \qquad (1.48б)$$

$$b_n=\frac{2}{T}\int\limits_{t_0}^{t_0+T} f(t)\sin n\,\omega_0 t dt. \qquad (1.48в)$$

Из ф-лы (1.48а) вытекает, что первый член ряда a_0 есть среднее значение $f(t)$ на интервале $(t_0,\ t_0+T)$. Таким образом, a_0 — постоянная составляющая функции $f(t)$ на этом интервале.

Тригонометрический ряд (1.46) можно записать как

$$f(t)=\sum_{n=0}^{\infty} C_n\cos(n\,\omega_0 t+\varphi_n), \qquad (1.49)$$

где

$$C_n=\sqrt{a_n^2+b_n^2},$$

$$\varphi_n=-\operatorname{arc\,tg}(b_n/a_n). \qquad (1.50)$$

Пример 1.4

Разложим функцию $f(t)$ (рис. 1.8а) в тригонометрический ряд Фурье на интервале $(0,1)$.

Очевидно,

$$f(t)=A\,t\,(0<t<1),\quad T=1 \quad \text{и} \quad \omega_0=2\,\pi/T=2\,\pi.$$

Следует выбрать $t_0=0$. Тогда

$$f(t)=a_0+a_1\cos 2\,\pi\, t+a_2\cos 4\,\pi\, t+\ldots+a_n\cos 2\,\pi\, nt+$$
$$+\,b_n\sin 2\,\pi\, t+b_2\sin 4\,\pi\, t+\ldots+b_n\sin 2\,\pi\, nt+\ldots \qquad (1.51)$$

Коэффициенты ряда (1.51) определяются по ф-лам (1.48а)—(1.48в):

$$a_0=\frac{1}{T}\int\limits_0^T f(t)\,dt=1\cdot\int\limits_0^1 A\,t\,dt=\frac{A}{2}. \qquad (1.52а)$$

Аналогично

$$a_n=\frac{2}{1}\int_0^1 A\,t\cos 2\pi n\,t\,dt=\frac{A}{2\pi^2 n^2}[\cos 2\pi nt+2\pi n\,t\sin 2\pi n\,t]_0^1=0, \qquad (1.52б)$$

$$b_n=\frac{2}{1}\int_0^1 At\sin 2\pi n\,t\,dt=\frac{A}{2\pi^2 n^2}[\sin 2\pi nt-2\pi nt\cos 2\pi nt]_0^1=-\frac{A}{\pi n}. \qquad (1.52в)$$

Так как $a_n=0$ для всех n, все косинусоидальные члены ряда (1.51) равны нулю. Коэффициенты при синусоидальных членах определяются выражением (1.52в). Окончательно ряд (1.51) принимает вид:

$$f(t)=\frac{A}{2}-\frac{A}{\pi}\sin 2\pi t-\frac{A}{2\pi}\sin 4\pi t-\frac{A}{3\pi}\sin 6\pi t-$$

$$-\frac{A}{n\pi}\sin 2\pi nt-\ldots=\frac{A}{2}-\frac{A}{\pi}\sum_{n=1}^{\infty}\frac{\sin 2\pi nt}{n} \quad (0<t<1) \qquad (1.53)$$

Заметим, что рядом (1.53) можно представить любую другую функцию, совпадающую с $f(t)$ на интервале (0,1). Например, функции $f_1(t)$ и $f_2(t)$ на рис. 1.8б и 1.8в совпадают с $f(t)$ на интервале (0,1), и поэтому обе эти функции на интервале (0,1) представляются также рядом (1.53).

ЭКСПОНЕНЦИАЛЬНЫЙ РЯД ФУРЬЕ

Покажем, что множество комплексных экспоненциальных функций $\{e^{in\omega_0 t}\}$ $(n=0, \pm 1, \pm 2...)$ ортогонально на интервале $(t_0, t_0+\frac{2\pi}{\omega_0})$ при любом t_0. Ортогональность этих функций можно доказать, рассмотрев интеграл

$$I=\int_{t_0}^{t_0+\frac{2\pi}{\omega_0}}(e^{in\omega_0 t})(e^{im\omega_0 t})^*\,dt=\int_{t_0}^{t_0+\frac{2\pi}{\omega_0}}e^{in\omega_0 t}e^{-im\omega_0 t}\,dt.$$

При $n=m$ $I=\int_{t_0}^{t_0+\frac{2\pi}{\omega_0}}dt=\frac{2\pi}{\omega_0}$,

при $n\neq m$ $I=\frac{1}{i(n-m)\omega_0}e^{i(n-m)\omega_0 t_0}[e^{i2\pi(n-m)}-1]$.

Так как n и m — целые числа, то $e^{i2\pi(n-m)}=1$ и $I=0$. Таким образом,

$$\int_{t_0}^{t_0+\frac{2\pi}{\omega_0}}e^{in\omega_0 t}(e^{im\omega_0 t})^*\,dt=\begin{cases}\frac{2\pi}{\omega_0} & \text{при } m=n\\ 0 & \text{при } m\neq n.\end{cases} \qquad (1.54)$$

Как и ранее, положим $T=2\pi/\omega_0$. Из ф-лы (1.54) следует, что множество функций $\{e^{in\omega_0 t}\}$ $(n=0, \pm 1, \pm 2...)$ ортогонально на

интервале (t_0, t_0+T). Более того, можно показать, что это множество является полным. Поэтому произвольную функцию $f(t)$ можно представить на интервале (t_0, t_0+T) линейной комбинацией экспоненциальных функций:

$$f(t)=F_0+F_1 e^{i\omega_0 t}+F_2 e^{i2\omega_0 t}+\ldots+F_n e^{in\omega_0 t}+\ldots+$$
$$+F_{-1}e^{-i\omega_0 t}+F_{-2}e^{-i2\omega_0 t}+\ldots+F_{-n}e^{-in\omega_0 t}+\ldots=$$
$$=\sum_{n=-\infty}^{\infty} F_n e^{in\omega_0 t} \quad (t_0<t<t_0+T), \tag{1.55}$$

где $\omega_0=2\pi/T$.

Суммирование в ф-ле (1.55) выполняется по целым n от $-\infty$ до ∞, включая нуль. Представление $f(t)$ рядом (1.55) известно как экспоненциальный ряд Фурье. Коэффициенты этого ряда определяются по ф-ле (1.42):

$$F_n=\frac{\int\limits_{t_0}^{t_0+T} f(t)\,(e^{in\omega_0 t})^*\,dt}{\int\limits_{t_0}^{t_0+T} e^{in\omega_0 t}\,(e^{in\omega_0 t})^* dt}=\frac{\int\limits_{t_0}^{t_0+T} f(t)\,e^{-in\omega_0 t}\,dt}{\int\limits_{t_0}^{t_0+T} e^{in\omega_0 t}\,e^{-in\omega_0 t}\,dt}=\frac{1}{T}\int\limits_{t_0}^{t_0+T} f(t)\,e^{-in\omega_0 t}\,dt. \tag{1.56}$$

Этот результат можно получить и непосредственно, помножив обе части равенства (1.55) на $e^{-in\omega_0 t}$ и проинтегрировав их по t от t_0 до t_0+T. В силу ортогональности все члены равенства, кроме одного в правой части, равны нулю и выражение для F_n получается в виде (1.56).

Таким образом, любую функцию $f(t)$ на интервале $(t_0<t<t_0+T)$ можно выразить дискретной суммой экспоненциальных функций $\{e^{in\omega_0 t}\}$ $(\omega_0=2\pi/T,\ n=0,\ \pm1,\ \pm2\ldots)$ [ф-лы (1.55), (1.56)]. Следует отметить, что тригонометрический и экспоненциальный ряды Фурье не являются двумя различными типами рядов, а представляют собой два различных способа выражения одного ряда. Коэффициенты одного ряда можно выразить через коэффициенты другого, в чем нетрудно убедиться при помощи ф-л (1.48) и (1.56), из которых вытекает, что

$$a_0=F_0,\quad a_n=F_n+F_{-n},\quad b_n=i\,(F_n-F_{-n}) \tag{1.57}$$

и

$$F_n=\frac{1}{2}(a_n-i\,b_n). \tag{1.58}$$

Для иллюстрации изложенного рассмотрим функцию $f(t)$ из примера 1.4. Эта функция была представлена тригонометрическим (1.53) и экспоненциальным (1.55) рядами Фурье. Воспользуемся ф-лой (1.58) для получения коэффициентов экспоненциального ря-

да из коэффициентов тригонометрического. Подставив выражения (1.52) в ф-лу (1.58), получим

$$F_0 = A/2, \quad F_n = \mathrm{i}\, A/2\pi n. \tag{1.59}$$

Следовательно,

$$f(t) = \frac{A}{2} + \frac{\mathrm{i}\, A}{2\pi} \sum_{n=-\infty}^{\infty} \frac{1}{n}\, \mathrm{e}^{\mathrm{i}2\pi n t}\,. \tag{1.60}$$

1.4. Представление произвольной периодической функции рядом Фурье на бесконечном интервале $(-\infty < t < \infty)$

До сих пор функция $f(t)$ представлялась рядом Фурье на конечном интервале $(t_0,\ t_0+T)$. Вне этого интервала функция $f(t)$ и соответствующий ей ряд Фурье могут не совпадать. Если же функция $f(t)$ периодическая, можно показать, что представление применимо ко всему бесконечному интервалу $(-\infty,\ \infty)$. В этом легко убедиться, рассмотрев некоторую функцию $f(t)$ и соответствующий ей экспоненциальный ряд Фурье на интервале $(t_0,\ t_0+T)$ [см. ф-лу (1.55)]:

$$f(t) = \sum_{n=-\infty}^{\infty} F_n\, \mathrm{e}^{\mathrm{i} n\omega_0 t}\,.$$

Равенство справедливо на интервале $(t_0 < t < t_0+T)$. Вне указанного интервала левая и правая части равенства могут быть не равны. Однако легко видеть, что правая часть этого равенства является периодической функцией (с периодом $T = 2\pi/\omega_0$), так как

$$\mathrm{e}^{\mathrm{i}\, n\omega_0 t} = \mathrm{e}^{\mathrm{i}\, n\omega_0 (t+T)} \tag{1.61}$$

Поэтому, если $f(t)$ — периодическая функция с периодом T, то равенство (1.55) справедливо для всего интервала $(-\infty,\ \infty)$. Заметим, что выбор величины t_0 несуществен.

Пример 1.5

Разложим в экспоненциальный ряд Фурье «выпрямленное» синусоидальное колебание, изображенное на рис. 1.9.

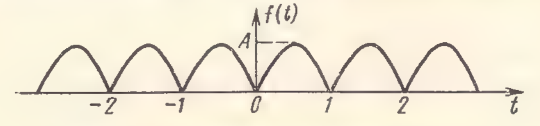


Рис. 1.9. Выпрямленная синусоидальная функция

Для этой функции

$$f(t) = \sum_{n=-\infty}^{\infty} F_n\, \mathrm{e}^{\mathrm{i}\, n\, \omega_0\, t} \qquad \begin{cases} \omega_0 = 2\pi, \\ T = 1, \end{cases}$$

$$F_n = A \int_0^1 \sin \pi t \, e^{-i 2 \pi n t} \, dt = \frac{-2A}{\pi (4n^2 - 1)}$$

Следовательно,

$$f(t) = \frac{-2A}{\pi} \sum_{n=-\infty}^{\infty} \frac{1}{4n^2 - 1} e^{i 2 \pi n t}. \qquad (1.62)$$

1.5. Комплексный спектр сигнала

Разложение в ряд Фурье периодической функции с периодом T показывает, что она имеет частотные составляющие с угловыми частотами ω_0, $2\omega_0$, $3\omega_0, \ldots, n\omega_0 \ldots$, где $\omega_0 = 2\pi/T$. Таким образом, периодическая функция $f(t)$ обладает своим спектром частот. Если функция $f(t)$ известна, можно определить ее спектр и, наоборот, по известному спектру можно найти соответствующую периодическую функцию $f(t)$. Следовательно, возможны два представления периодической функции $f(t)$: *временно́е*, при котором $f(t)$ выражается как функция времени, и *частотное*, при котором определен спектр (т. е. амплитуды различных частотных составляющих). Заметим, что спектр существует только на частотах ω_0, $2\omega_0$, $3\omega_0 \ldots$ и т. д., т. е. его нельзя представить непрерывной кривой. Это *дискретный спектр*, который иногда называют *линейчатым*. Такой спектр можно изобразить графически в виде вертикальных линий на частотах $\omega = \omega_0$, $2\omega_0 \ldots$ и т. д., причем высота каждой линии пропорциональна амплитуде соответствующей частотной составляющей.

Для представления спектра можно использовать тригонометрический или экспоненциальный ряды. Для рассматриваемых здесь задач предпочтительнее экспоненциальная форма ряда, когда периодическая функция выражается суммой экспоненциальных функций с частотами 0, $\pm\omega_0$, $\pm 2\omega_0 \ldots$ и т. д. Нетрудно понять значение отрицательных частот. Оба сигнала $e^{i\omega t}$ и $e^{-i\omega t}$ изменяются с одинаковой частотой ω. Их можно представить двумя векторами, вращающимися в противоположных направлениях и при сложении дающих действительную функцию времени $e^{i\omega t} + e^{-i\omega t} = 2\cos\omega t$. Для периодической функции с периодом T экспоненциальный ряд имеет вид

$$f(t) = F_0 + F_1 e^{i\omega_0 t} + F_2 e^{i2\omega_0 t} + \ldots + F_n e^{in\omega_0 t} + \ldots +$$
$$+ F_{-1} e^{-i\omega_0 t} + F_{-2} e^{-i2\omega_0 t} + \ldots + F_{-n} e^{-in\omega_0 t} + \ldots$$

Обычно амплитуды F_n являются комплексными и характеризуются величиной и фазой. Поэтому в общем случае для частотного представления периодической функции необходимо иметь два линейчатых спектра: *спектр амплитуд и спектр фаз.* В большинстве случаев, однако, амплитуды частотных составляющих явля-

ются только действительными или только мнимыми, что дает возможность описывать функцию только одним спектром.

Рассмотрим периодическую функцию из примера 1.5 (см. рис. 1.9). Это «выпрямленная» синусоидальная функция, экспоненциальный ряд Фурье для которой был получен ранее [ф-ла (1.62)]:

$$f(t)=\frac{2A}{\pi}-\frac{2A}{3\pi}e^{i2\pi t}-\frac{2A}{15\pi}e^{i4\pi t}-\frac{2A}{35\pi}e^{i6\pi t}-\ldots-$$
$$-\frac{2A}{3\pi}e^{-i2\pi t}-\frac{2A}{15\pi}e^{-i4\pi t}-\frac{2A}{35\pi}e^{-i6\pi t}-\ldots$$

Спектр существует на частотах $\omega=0,\ \pm2\pi,\ \pm4\pi,\ \pm6\pi\ldots$ и т. д. с соответствующими амплитудами составляющих $2A/\pi$, $-2A/3\pi$, $-2A/15\pi$, $-2A/35\pi$ и т. д. Заметим, что все амплитуды являются действительными, и поэтому необходимо дать график только одного спектра (рис. 1.10). Спектр симметричен относительно вертикальной оси, проходящей через начало координат. Это не случайное совпадение. Покажем, что спектр амплитуд любой периодической функции симметричен относительно вертикальной оси, проходящей через начало координат. Коэффициенты разложения

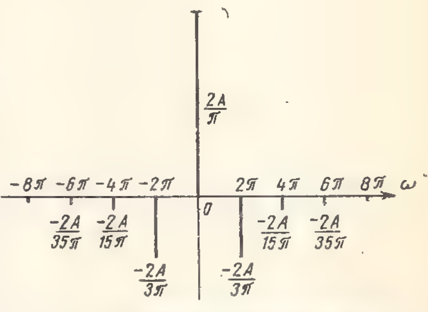


Рис. 1.10. Дискретный спектр выпрямленной синусоидальной функции

$$F_n=\frac{1}{T}\int\limits_{-T/2}^{T/2}f(t)\,e^{-in\omega_0 t}\,dt \quad\text{и}\quad F_{-n}=\frac{1}{T}\int\limits_{-T/2}^{T/2}f(t)\,e^{in\omega_0 t}\,dt.$$

Из этих выражений видно, что F_n и F_{-n} являются комплексно-сопряженными величинами, т. е. $F_{-n}=F_n^*$, следовательно, $|F_n|=|F_{-n}|$. Таким образом, спектр амплитуд симметричен относительно вертикальной оси, проходящей через начало координат, и является четной функцией ω. Если F_n — действительная величина, то F_{-n} — также действительная и $F_n=F_{-n}$. Если F_n — комплексная величина:

$$F_n=|F_n|\,e^{i\theta_n},$$

то

$$F_{-n}=|F_n|\,e^{-i\theta_n}. \tag{1.63}$$

Следовательно, спектр фаз — нечетная функция относительно вертикальной оси, проходящей через начало координат.

Пример 1.6

Разложим периодическую стробирующую функцию, показанную на рис. 1.11, в экспоненциальный ряд Фурье и изобразим ее частотный спектр.

Стробирующая функция имеет ширину (длительность импульса) δ и период повторения T, *сек.* На интервале в один период функцию можно аналитически

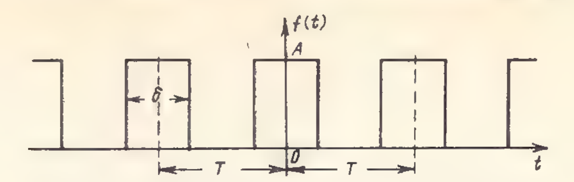


Рис. 1.11. Периодическая стробирующая функция

записать как

$$f(t)=\begin{cases} A \text{ при } \left(-\frac{\delta}{2}<t<\frac{\delta}{2}\right), \\ 0 \text{ при } \left(\frac{\delta}{2}<t<T-\frac{\delta}{2}\right). \end{cases}$$

Для удобства выберем пределы интегрирования от $-\frac{\delta}{2}$ до $T-\frac{\delta}{2}$:

$$F_n=\frac{1}{T}\int\limits_{-\frac{\delta}{2}}^{T-\frac{\delta}{2}} f(t)\,e^{-i n\omega_0 t}\,dt=\frac{1}{T}\int\limits_{-\frac{\delta}{2}}^{\frac{\delta}{2}} A\,e^{-i n\omega_0 t}\,dt=$$

$$=\frac{-A}{i n\omega_0 T}\,e^{-i n\omega_0 t}\Big|_{-\delta/2}^{\delta/2}=\frac{2A}{n\omega_0 T}\,\frac{e^{i n\omega_0\frac{\delta}{2}}-e^{-i n\omega_0\frac{\delta}{2}}}{2i}=$$

$$=\frac{2A}{n\omega_0 T}\sin\left(\frac{n\omega_0\delta}{2}\right)=\frac{A\delta}{T}\left[\frac{\sin\left(\frac{n\omega_0\delta}{2}\right)}{\frac{n\omega_0\delta}{2}}\right]. \qquad (1.64)$$

Заключенная в скобки функция имеет форму $(\sin x)/x$. Эта функция играет существенную роль в теории связи и называется *функцией отсчетов*[1]). В дальнейшем эта функция обозначается:

$$Sa(x)=(\sin x)/x. \qquad (1.65)$$

Функция отсчетов показана на рис. 1.12. Заметим, что она осциллирует с периодом 2π, спадая по амплитуде с увеличением x и пе-

[1]) Отсчетом (или выборочным значением) называется мгновенное значение сигнала в некоторый момент времени. Название «функция отсчетов» объясняется тем, что функция $(\sin x)/x$ используется для восстановления непрерывного сигнала по совокупности его мгновенных значений (см. § 1.15). — *Прим. ред.*

реходя через нуль в точках $x=\pm\pi$, $\pm 2\pi$, $\pm 3\pi$ и т. д. Из ф-лы (1.64) следует, что

$$F_n = \frac{A\,\delta}{T} Sa\left(\frac{n\,\omega_0\delta}{2}\right).$$

Так как $\omega_0 = 2\pi/T$, $n\omega_0\delta/2 = n\pi\delta/T$, получаем

$$F_n = \frac{A\,\delta}{T} Sa\left(\frac{n\,\pi\delta}{T}\right) \qquad (1.66а)$$

и

$$f(t) = \frac{A\,\delta}{T} \sum_{n=-\infty}^{\infty} Sa\left(\frac{n\,\pi\delta}{T}\right) e^{in\omega_0 t}. \qquad (1.66б)$$

Из ф-л (1.66) следует, что F_n — действительная величина, поэтому для частотного представления достаточен лишь один спектр,

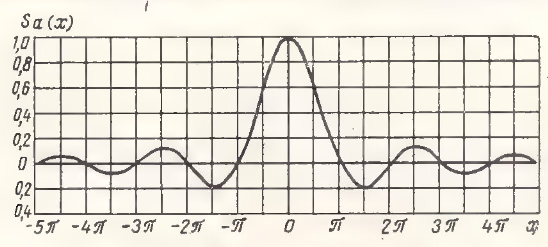


Рис. 1.12. Функция отсчетов $Sa(x)$

Кроме того, поскольку $Sa(x)$ — четная функция, то $F_n = F_{-n}$. Найденный частотный спектр является дискретной функцией, существующей только на частотах $\omega = 0$, $\pm 2\pi/T$, $\pm 4\pi/T$, $\pm 6\pi/T$ и т. д. с соответствующими амплитудами $A\delta/T$, $(A\delta/T)\,Sa\,(\pi\delta/T)$, $(A\delta/T)\,Sa(2\pi\delta/T)$ ... Рассмотрим зависимость спектра от величины δ и T. Положим длительность импульса равной 0,05 *сек*, а период T выберем равным 0,25 *сек*, 0,5 *сек* и 1 *сек*. Вычисленные по ф-ле (1.66а) спектры показаны на рис. 1.13.

Очевидно, с увеличением периода T основная частота $2\pi/T$ уменьшается и вследствие этого растет число частотных составляющих, приходящихся на некоторый интервал частот. Иными словами, с увеличением периода T спектр становится плотнее. Однако амплитуды частотных составляющих при этом уменьшаются. Устремляя T к бесконечности, получаем одиночный прямоугольный импульс длительностью δ с частотой повторения, равной нулю. Заметим, что форма частотного спектра сохраняется неизменной, т. е. огибающая спектра зависит только от формы импульса, но не от периода повторения T. В пределе при $T \to \infty$ функция $f(t)$ состоит из одного неповторяющегося импульса, и спектр представляет, таким образом, непериодическую функцию на бесконечном интервале $(-\infty < t < \infty)$.

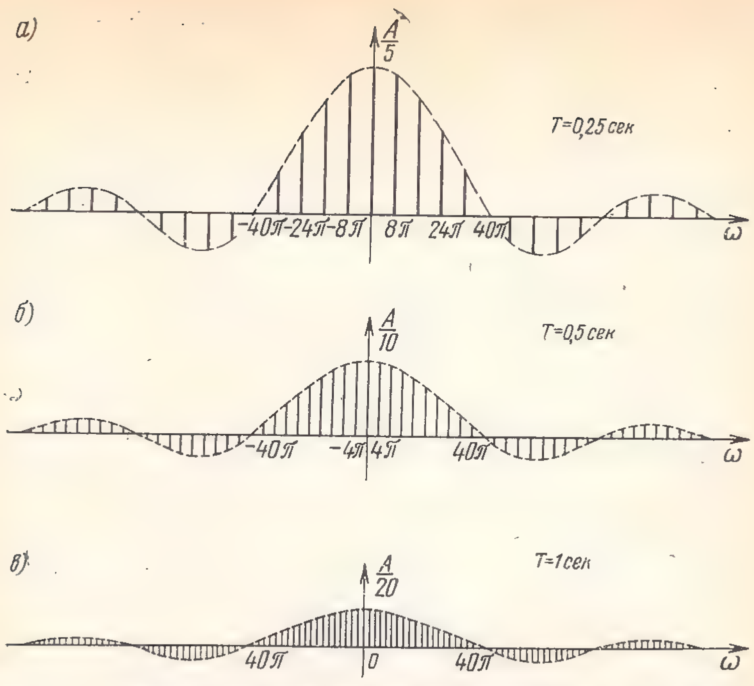


Рис. 1.13. Спектры периодической стробирующей функции при различных периодах повторения:
а) $\delta/T=1/5$, $\delta=1/20$, $T=1/4$; *б)* $\delta/T=1/10$, $\delta=1/20$, $T=1/2$;
в) $\delta/T=1/20$, $\delta=1/20$, $T=1$

Таким образом, представление периодической функции суммой экспоненциальных функций распространяется и на случай непериодической функции. Подробнее этот вопрос обсуждается в следующем параграфе.

1.6. Представление произвольной функции на бесконечном интервале $(-\infty, \infty)$. Преобразование Фурье

Как уже указывалось, произвольную функцию можно представить экспоненциальным (или тригонометрическим) рядом на конечном интервале. В частном случае периодической функции это представление можно распространить на бесконечный интервал $(-\infty, \infty)$. Желательно, однако, уметь представлять любые (периодические или непериодические) сигналы на бесконечном интервале $(-\infty, \infty)$ с помощью экспоненциальных функций. Покажем, что непериодический сигнал можно выразить непрерывной суммой (интегралом) экспоненциальных функций, в отличие от периодических сигналов, которые представляются дискретной суммой экспоненциальных функций.

Поставленную задачу можно решить двумя способами. Первый из них заключается в выражении сигнала $f(t)$ через экспоненциальные функции на конечном интервале $(-T/2<T<T/2)$, а затем период T устремляется к бесконечности. Второй способ сводится к созданию периодической функции с периодом T, которая совпадает с $f(t)$ только в пределах одного периода. В пределе при $T \to \infty$ оказывается, что периодическая функция имеет один единственный период на интервале $(-\infty<t<\infty)$, что и соответствует функции $f(t)$. Между указанными вариантами решения нет существенного различия, однако последний вариант более удобен, так как позволяет проследить процесс предельного перехода без изменения формы частотного спектра. Такой предельный переход отчасти уже рассматривался в примере 1.6 для случая периодической стробирующей функции (см. рис. 1.11). Было выяснено, что с увеличением периода основная частота уменьшается и частотный спектр становится плотнее, т. е. на данном частотном интервале располагается больше частотных составляющих. При этом амплитуды составляющих уменьшаются. Однако форма частотного спектра остается неизменной, что легко заметить из графиков на рис. 1.13.

Пусть задана функция $f(t)$, показанная на рис. 1.14*а*, которую необходимо представить на бесконечном интервале $(-\infty<t<\infty)$

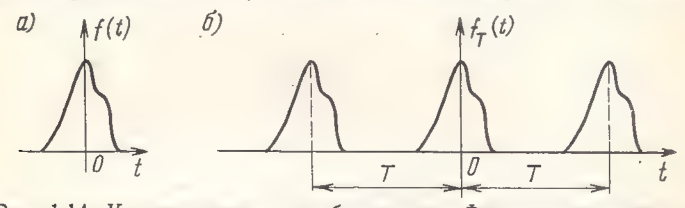


Рис. 1.14. К определению преобразования Фурье непериодической функции:
а) непериодическая функция $f(t)$; *б)* периодическая функция, образованная повторением $f(t)$

суммой экспоненциальных функций. С этой целью построим новую периодическую функцию $f_T(t)$ с периодом T, в которой функция $f(t)$ повторяется через каждые T *сек* (рис. 1.14*б*). Период T выбираем достаточно большим, чтобы соседние импульсы $f(t)$ не перекрывались. Эта новая функция $f_T(t)$ является периодической и, следовательно, ее можно представить экспоненциальным рядом Фурье. В пределе при $T \to \infty$ функция $f_T(t)$ переходит в $f(t)$: $\lim\limits_{T\to\infty} f_T(t) = f(t)$. Таким образом, ряд Фурье, представляющий $f_T(t)$ на бесконечном интервале, будет также представлять $f(t)$ на том же интервале, если в выражении ряда положить $T=\infty$.

Экспоненциальный ряд Фурье для $f_T(t)$ имеет вид

$$f_T(t) = \sum_{n=-\infty}^{\infty} F_n e^{jn\omega_0 t},$$

где

$$F_n = \frac{1}{T} \int\limits_{-T/2}^{T/2} f_T(t) e^{-i n\omega_0 t}\, dt. \qquad (1.67)$$

Коэффициент F_n характеризует амплитуду составляющей на частоте $n\omega_0$. Пусть теперь T будет достаточно большим. Чем больше становится T, тем меньше ω_0 (основная частота) и спектр оказывается плотнее. Как видно из ф-лы (1.67), амплитуды отдельных спектральных составляющих при этом уменьшаются, но форма частотного спектра остается неизменной. В пределе при $T \to \infty$ амплитуды частотных составляющих становятся бесконечно малыми, но при этом число составляющих стремится к бесконечности. Теперь спектр существует на любой частоте ω и из дискретной функции частоты превращается в непрерывную. Введем новые обозначения:

$$n\,\omega_0 = \omega_n. \qquad (1.68)$$

Так как F_n является функцией ω_n, заменим обозначение на $F_n(\omega_n)$. Кроме того, обозначим

$$TF_n(\omega_n) = F(\omega_n). \qquad (1.69)$$

Теперь

$$f_T(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega_n) e^{i\omega_n t}, \qquad (1.70)$$

$$F(\omega_n) = TF_n = \int\limits_{-T/2}^{T/2} f_T(t) e^{-i\omega_n t}\, dt. \qquad (1.71)$$

Подставив в ф-лу (1.70) значение $T = 2\pi/\omega_0$, получим

$$f_T(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} F(\omega_n)\, e^{i\omega_n t}\, \omega_0. \qquad (1.72)$$

Равенство (1.72) показывает, что $f_T(t)$ можно выразить суммой экспоненциальных функций с частотами $\omega_1, \omega_2, \ldots, \omega_n$. Амплитуда составляющей на частоте ω_n равна $F(\omega_n)\omega_0/2\pi$, т. е. не равна $F(\omega_n)$, но пропорциональна ей.

Попытаемся графически проиллюстрировать ф-лу (1.72), которая является дискретной суммой или суммой дискретных частотных составляющих. В действительности величина $F(\omega_n)e^{i\omega_n t}$ в общем случае является комплексной и графическое представление должно состоять из двух графиков (действительной и мнимой частей или же графиков амплитудного и фазового спектров). Предположим, однако, что величина $F(\omega_n)e^{i\omega_n t}$ — действительная (в дальнейшем будет показано, что такое предположение допустимо).

На рис. 1.15 показан график этой величины в зависимости от ω. Функция существует только при дискретных значениях $\omega=\omega_1, \omega_2, \ldots, \omega_n$, где $\omega_n = n\omega_0$.

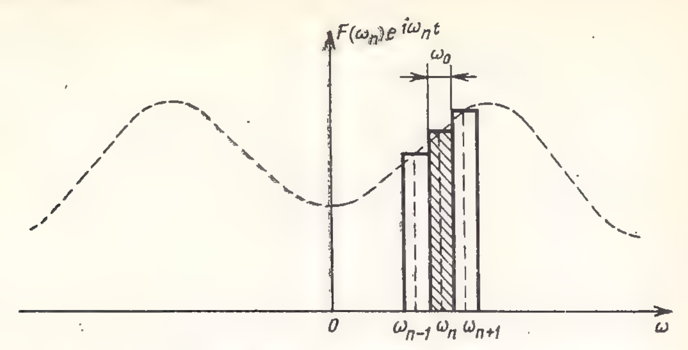


Рис. 1.15. Графическое пояснение интеграла Фурье

Интервал между соседними частотными составляющими равен ω_0, поэтому площадь заштрихованного прямоугольника на графике 1.15, очевидно, равна $F(\omega_n)e^{in\omega_n t}\omega_0$. Формула (1.72) представляет собой сумму площадей таких прямоугольников для значений n от $-\infty$ до ∞. Приближенно сумма площадей прямоугольников равна площади под пунктирной кривой. Чем меньше ω_0, тем лучше аппроксимация. В пределе при $T\to\infty$ величина ω_0 становится бесконечно малой и ее можно обозначить как $d\omega$. Дискретная сумма в ф-ле (1.72) переходит в интеграл, что соответствует площади под этой кривой. Кривая теперь оказывается непрерывной функцией частоты и записывается как $F(\omega)e^{i\omega t}$. При $T\to\infty$ функция $f_T(t)\to f(t)$ и ф-лы (1.70) и (1.71) принимают вид:

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}F(\omega)e^{i\omega t}\,d\omega, \qquad (1.73)$$

$$F(\omega)=\int_{-\infty}^{\infty}f(t)\,e^{-i\omega t}\,dt. \qquad (1.74)$$

Равенство (1.73) представляет непериодическую функцию $f(t)$ как непрерывную сумму экспоненциальных функций с частотами в интервале $(-\infty<\omega<\infty)$. Амплитуда составляющей на любой частоте ω пропорциональна $F(\omega)$, поэтому $F(\omega)$ является частотным спектром функции $f(t)$ и называется *функцией спектральной плотности.* Заметим, что частотный спектр непериодического сигнала является непрерывной функцией частоты. Спектральная плотность $F(\omega)$ находится по ф-ле (1.74).

Формулы (1.73) и (1.74) известны как пара преобразований Фурье: ф-ла (1.74) называется прямым преобразованием Фурье

сигнала $f(t)$, а ф-ла (1.73) — обратным преобразованием функции $F(\omega)$. Символически эти преобразования записываются в виде

$$F(\omega)=\mathcal{F}[f(t)], \quad f(t)=\mathcal{F}^{-1}[F(\omega)]. \tag{1.75}$$

1.7. Некоторые замечания о непрерывной спектральной функции

Выше было показано, что непериодический сигнал $f(t)$ можно представить непрерывной суммой экспоненциальных функций, частоты которых лежат в интервале $(-\infty, \infty)$. Амплитуды составляющих при этом бесконечно малы и пропорциональны спектральной плотности $F(\omega)$.

Понятие непрерывного спектра вызывает иногда затруднения, так как для периодической функции спектр существует на дискретных частотах с конечными амплитудами. Однако оно становится более доступным, если рассмотреть аналогичное конкретное явление. Одним из примеров непрерывного распределения является нагруженная балка. Рассмотрим балку с действующими на нее грузами $F_1, F_2, F_3, \ldots, F_7$, равномерно приложенными в точках $x_1, x_2, x_3, \ldots, x_7$ (рис. 1.16*а*). Балка нагружена в семи дискретных точках и общий действующий на нее груз определяется суммой этих отдельных нагрузок: $W_T=\sum_{r=1}^{7} F_r$.

Теперь рассмотрим случай непрерывно нагруженной балки, показанной на рис. 1.16*б*. Плотность нагрузки $F(x)$, *кг/м* является функцией x. Полная нагруз-

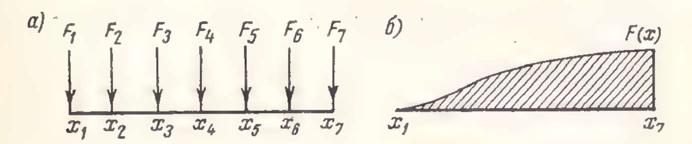


Рис. 1.16. Нагруженная балка:
а) нагрузка приложена в дискретных точках; *б)* нагрузка непрерывно распределена

ка, приходящаяся на балку, определяется теперь непрерывной суммой нагрузок, т. е. интегралом от $F(x)$ по всей длине балки: $W_T=\int_{x_1}^{x_2} F(x)dx$.

В случае дискретной нагрузки груз находится только в дискретных точках. В других точках нагрузки нет. С другой стороны, в случае непрерывно распределенной нагрузки последняя существует в каждой точке, но в какой-либо определенной точке она равна нулю. На небольшом интервале dx нагрузка равна $F(x)dx$, поэтому $F(x)$ представляет относительную величину нагрузки на интервале dx вблизи точки x. Аналогичное явление имеет место в случае сигналов и их частотных спектров. Периодический сигнал можно представить суммой дискретных экспоненциальных сигналов с конечными амплитудами

$$f(t)=\sum_{n=-\infty}^{\infty} F_n e^{j\omega_n t} \quad (\omega_n=n\omega).$$

Для непериодического сигнала распределение амплитуд экспоненциальных функций становится непрерывным, т. е. спектральная функция существует на всех частотах ω, но амплитуда частотной составляющей на каждой конкретной частоте равна нулю. Каждое бесконечно малое слагаемое имеет амплитуду

$(1/2\pi)\ F(\omega)d\omega$, а функция $f(t)$ выражается непрерывной суммой таких бесконечно малых составляющих [ф-ла (1.73)].

Коэффициент $1/2\pi$ в ф-ле (1.73) можно убрать, если переменную интегрирования ω заменить на f. Тогда $\omega=2\pi f$, $d\omega=2\pi df$ и выражение (1.73) примет вид

$$f(t)=\int_{-\infty}^{\infty} F(2\pi f)\,e^{i2\pi ft}\,df. \qquad (1.76)$$

1.8. Временнóе и частотное представления сигналов

Преобразование Фурье служит инструментом, позволяющим представлять данный сигнал экспоненциальными составляющими. Функция $F(\omega)$ есть прямое преобразование Фурье сигнала $f(t)$; она характеризует относительные амплитуды различных частотных составляющих. Поэтому $F(\omega)$ представляет сигнал $f(t)$ в частотной области. Временнóе представление определяет некоторый сигнал в каждый момент времени, тогда как частотное представление характеризует относительные амплитуды частотных составляющих сигнала. Любое из этих представлений полностью определяет сигнал.

Однако функция $F(\omega)$ в общем случае комплексна:

$$F(\omega)=|F(\omega)|\,e^{i\theta(\omega)}$$

и для ее представления необходимы два графика: амплитудного $|F(\omega)|$ и фазового $\theta(\omega)$ спектров. Во многих случаях $F(\omega)$ — либо действительная, либо мнимая функция, и поэтому для ее представления достаточно одного графика. Если $f(t)$ — действительная функция, то согласно ф-ле (1.74)

$$F(-\omega)=\int_{-\infty}^{\infty} f(t)\,e^{i\omega t}\,dt=F^{*}(\omega). \qquad (1.77)$$

Таким образом, если $F(\omega)=|F(\omega)|e^{i\theta(\omega)}$, то

$$F(-\omega)=|F(\omega)|\,e^{-i\theta(\omega)}. \qquad (1.78)$$

Из этих равенств очевидно, что амплитудный спектр $|F(\omega)|$ является четной функцией ω, а фазовый спектр $\theta(\omega)$ нечетной.

1.9. Существование преобразования Фурье

Из формулы преобразования Фурье (1.74) следует, что если интеграл $\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt$ имеет конечное значение, то преобразование Фурье существует. Так как модуль $e^{-i\omega t}$ равен единице, условие существования преобразования Фурье функции $f(t)$ состоит

в том, что интеграл $\int_{-\infty}^{\infty} |f(t)|dt$ должен иметь конечное значение. Однако если рассматриваются сингулярные функции (см. § 1.11), условие абсолютной интегрируемости не всегда необходимо. Ниже будет показано, что есть функции не абсолютно интегрируемые, но имеющие преобразование Фурье. Таким образом, абсолютная интегрируемость $f(t)$ является достаточным, но не необходимым условием для существования преобразования Фурье сигнала $f(t)$.

Такие функции, как $\sin\omega t$, $\cos\omega t$, единичный скачок $u_1(t)$ и т. д. не удовлетворяют упомянутому условию и, строго говоря, не имеют преобразования Фурье. Однако в пределе и для этих функций преобразование Фурье существует. Например, можно предположить, что функция $\sin\omega t$ существует только на интервале $-T/2 < t < T/2$. Такая функция имеет преобразование Фурье, пока T конечно. В пределе можно сделать T очень большим, но конечным. Ниже будет показано, что в пределе для таких функций преобразование Фурье существует

1.10. Преобразование Фурье некоторых функций

ОДНОСТОРОННИЙ ЭКСПОНЕНЦИАЛЬНЫЙ СИГНАЛ

$$f(t) = e^{-at} u_1(t),$$

$$\text{где} \quad u_1(t) = \begin{cases} 1 \text{ при } t > 0, \\ 0 \text{ при } t < 0, \end{cases}$$

$$F(\omega) = \int_{-\infty}^{\infty} e^{-at} u_1(t)\, e^{-i\omega t}\, dt = \int_{0}^{\infty} e^{-(a+i\omega)}\, dt = \frac{1}{a+i\omega} = \frac{1}{\sqrt{a^2+\omega^2}} e^{-i\,\mathrm{arctg}(\omega/a)} \tag{1.79}$$

Здесь

$$|F(\omega)| = 1/\sqrt{a^2+\omega^2} \text{ и } \theta(\omega) = -\mathrm{arc\,tg}\,(\omega/a). \tag{1.80}$$

Графики амплитудного $F(\omega)$ и фазового $\theta(\omega)$ спектров приведены в табл. 1.1 (поз. 1) [1]).

Заметим, что интеграл (1.79) сходится только при $a>0$. При $a<0$ преобразование Фурье не существует. Это подтверждается также и тем, что при $a<0$ функция $f(t)$ оказывается не абсолютно интегрируемой.

---

[1]) Заметим, что для большинства сигналов, приведенных в таблице, $F(\omega)$ является действительной функцией, и поэтому для ее представления достаточно только одного графика.

Таблица 1.1

НЕКОТОРЫЕ СИГНАЛЫ И ИХ ПРЕОБРАЗОВАНИЯ ФУРЬЕ

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
1	$f(t) = e^{-at} u_1(t)$	$f(t)$; $f(t)=e^{-at}u_1(t)$; 0; t	$F(\omega) = \dfrac{1}{a + i\omega}$	$\frac{1}{a}$; $\lvert F(\omega)\rvert$; $\frac{\pi}{2}$; 0; ω; $\theta(\omega)$; $-\frac{\pi}{2}$
2	$f(t) = t e^{-at} u_1(t)$	$f(t)$; $f(t)=te^{-at}u_1(t)$; 0; t	$F(\omega) = \dfrac{1}{(a + i\omega)^2}$	$\frac{1}{a^2}$; $\lvert F(\omega)\rvert$; π; $\theta(\omega)$; 0; ω; $-\pi$
3	$f(t) = e^{-a\lvert t\rvert}$	$f(t)$; $f(t)=e^{-a\lvert t\rvert}$; 0; t	$F(\omega) = \dfrac{2a}{a^2 + \omega^2}$	$\frac{2}{a}$; $F(\omega)$; $\theta(\omega)=0$; 0; ω

Продолжение табл. 1.1

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
4	$f(t) = e^{-at} \sin \omega_0 t \, u_1(t)$		$F(\omega) = \dfrac{\omega_0^2}{(a + i\omega)^2 + \omega_0^2}$	
5	$f(t) = Sa \dfrac{Wt}{2}$		$F(\omega) = \dfrac{\pi}{W} G_W(\omega) = \begin{cases} \dfrac{\pi}{W} & \text{при } \lvert\omega\rvert < W \\ 0 & \text{при } \lvert\omega\rvert > W \end{cases}$	
6	$f(t) = G_\tau(t) = \begin{cases} 1 & \text{при } \lvert t\rvert < \dfrac{\tau}{2}, \\ 0 & \text{при } \lvert t\rvert > \dfrac{\tau}{2} \end{cases}$		$F(\omega) = \tau \, Sa\left(\dfrac{\omega\tau}{2}\right)$	

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
7	$f(t)=\begin{cases}1-\dfrac{\lvert t\rvert}{\tau} & \text{при } \lvert t\rvert<\tau,\\ 0 & \text{при } \lvert t\rvert>\tau\end{cases}$	$f(t)$; 1; τ; 0; τ; t	$F(\omega)=\tau\left[Sa\left(\dfrac{\omega\tau}{2}\right)\right]^2$	$F(\omega)$; τ; $-\frac{8\pi}{\tau}$ $-\frac{6\pi}{\tau}$ $-\frac{4\pi}{\tau}$ $-\frac{2\pi}{\tau}$ 0 $\frac{2\pi}{\tau}$ $\frac{4\pi}{\tau}$ $\frac{6\pi}{\tau}$ $\frac{8\pi}{\tau}$ ω
8	$f(t)=e^{-t^2/(2\sigma^2)}$	$f(t)$; 1; $f(t)=e^{-t^2/2\sigma^2}$; 0; t	$F(\omega)=\sigma\sqrt{2\pi}\,e^{-\sigma^2\omega^2/2}$	$F(\omega)$; $\sigma\sqrt{2\pi}$; $\sigma\sqrt{2\pi}\,e^{-\sigma^2\omega^2/2}$; 0; ω
9	$f(t)=\delta(t)$	$f(t)$; $f(t)=\delta(t)$; 0; t	$F(\omega)=1$	$F(\omega)$; 1; 0; ω

Продолжение табл. 1.1

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
10	$f(t) = 1$		$F(\omega) = 2\pi\,\delta(\omega)$	
11	$f(t) = \cos\omega_0 t$		$F(\omega) = \pi\,[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$	
12	$f(t) = \sin\omega_0 t$		$F(\omega) = i\,\pi\,[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$	

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
13	$f(t) = \cos\omega_0 t \; u_1(t)$	$f(t)$; $f(t)=\cos\omega_0 t\, u(t)$; 0; t	$F(\omega) = \dfrac{\pi}{2}\,[\delta(\omega-\omega_0) + \delta(\omega+\omega_0)] + \dfrac{i\omega}{\omega_0^2-\omega^2}$	$\lvert F(\omega)\rvert$; $-\omega_0$; 0; ω_0; ω
14	$f(t) = \sin\omega_0 t \; u_1(t)$	$f(t)$; $f(t)=\sin\omega_0 t\, u(t)$; 0; t	$F(\omega) = \dfrac{\pi}{2i}\,[\delta(\omega-\omega_0) - \delta(\omega+\omega_0)] + \dfrac{\omega_0}{\omega_0^2-\omega^2}$	$\lvert F(\omega)\rvert$; $-\omega_0$; 0; ω_0; ω
15	$f(t) = u_1(t) = \begin{cases} 1 \text{ при } t>0, \\ 0 \text{ при } t<0 \end{cases}$	$f(t)$; $f(t)=u_1(t)$; 1; 0; t	$F(\omega) = \pi\,\delta(\omega) + \dfrac{1}{i\omega}$	$\lvert F(\omega)\rvert$; 0; ω

Продолжение табл. 1.1

№ пп.	Сигнал $f(t)$		Преобразование Фурье $F(\omega)$	
16	$f(t)=\delta_T(t)=$ $=\sum\limits_{n=-\infty}^{\infty}\delta(t-nT)$		$F(\omega)=\omega_0\,\delta_{\omega_0}(\omega)=$ $=\omega_0\sum\limits_{k=-\infty}^{\infty}\delta(\omega-k\,\omega_0)$ $(\omega_0=2\pi/T)$	
17	$f(t)=\vert t\vert$		$F(\omega)=-\dfrac{2}{\omega^2}$	
18	$f(t)=\operatorname{sign}(t)=$ $=\begin{cases}1 & \text{при } t>0\\ -1 & \text{при } t<0\end{cases}$		$F(\omega)=\dfrac{2}{i\,\omega}$	

ДВУСТОРОННИЙ ЭКСПОНЕНЦИАЛЬНЫЙ СИГНАЛ

$$f(t) = e^{-a|t|}$$

$$F(\omega) = \int_{-\infty}^{\infty} e^{-a|t|} e^{-i\omega t} dt = \int_{-\infty}^{0} e^{(a-i\omega)t} dt + \int_{0}^{\infty} e^{-(a+i\omega)t} = \frac{2a}{a^2 + \omega^2}. \quad (1.81)$$

Заметим, что в этом случае фазовый спектр $\theta(\omega) = 0$. График амплитудного спектра сигнала приведен в табл. 1.1 (поз. *3*).

СТРОБИРУЮЩАЯ ФУНКЦИЯ

Стробирующая функция $G_\tau(t)$ определяется как прямоугольный импульс и записывается в виде

$$G_\tau(t) = \begin{cases} 1 \text{ при } |t| < \tau/2, \\ 0 \quad\quad |t| > \tau/2. \end{cases}$$

Преобразование Фурье этой функции имеет вид

$$F(\omega) = \int_{-\tau/2}^{\tau/2} A e^{-i\omega\tau} dt = \frac{A}{i\omega} \left(e^{i\omega\tau/2} - e^{-i\omega\tau/2}\right) = A\tau \frac{\sin\frac{\omega\tau}{2}}{\omega\tau/2} = A\tau Sa(\omega\tau/2) \quad (1.82)$$

Заметим, что $F(\omega)$ — действительная функция и, следовательно, может быть представлена одним графиком (табл. 1.1, поз. *6*).

1.11. Сингулярные функции

Рассмотрим единичный скачок напряжения[1]), приложенного к конденсатору (рис. 1.17*а*). Ток через конденсатор $i = C\frac{du_1}{dt}$. Нетрудно видеть, что производная du_1/dt равна нулю для всех значений t, за исключением $t=0$, где она имеет неопределенное значение. В точке $t=0$ производная не существует, так как в этой точке функция $u_1(t)$ терпит разрыв. Здесь возникает серьезная математическая трудность, являющаяся следствием идеализации источника напряжения и элемента цепи. Решение существовало бы, если бы источник или конденсатор был неидеальным. Так, в случае источника с напряжением, показанным на рис. 1.18*а*, ток имеет форму прямоугольного импульса (рис. 1.18*б*). Решения для идеального единичного скачка напряжения не существует, но можно получить решение в пределе, рассмотрев неидеальный случай, изображенный на рис. 1.18*а*, а затем устремив a к нулю.

---

[1]) Эта функция называется иногда *единичной* и *функцией включения*. — *Прим. ред.*

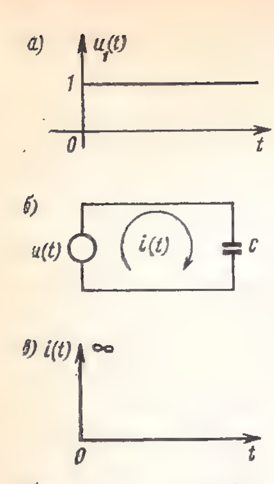
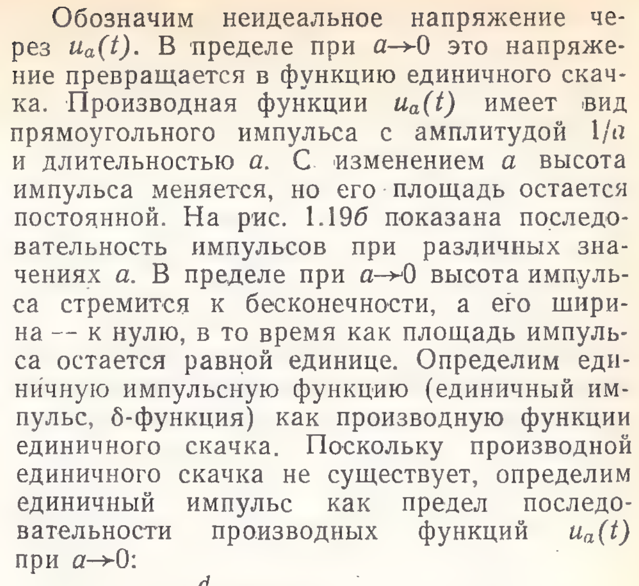
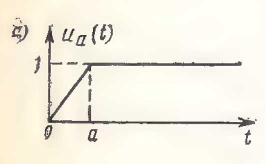


Рис. 1.17. Воздействие единичного скачка напряжения на конденсатор



Обозначим неидеальное напряжение через $u_a(t)$. В пределе при $a \to 0$ это напряжение превращается в функцию единичного скачка. Производная функции $u_a(t)$ имеет вид прямоугольного импульса с амплитудой $1/a$ и длительностью a. С изменением a высота импульса меняется, но его площадь остается постоянной. На рис. 1.19б показана последовательность импульсов при различных значениях a. В пределе при $a \to 0$ высота импульса стремится к бесконечности, а его ширина — к нулю, в то время как площадь импульса остается равной единице. Определим единичную импульсную функцию (единичный импульс, δ-функция) как производную функции единичного скачка. Поскольку производной единичного скачка не существует, определим единичный импульс как предел последовательности производных функций $u_a(t)$ при $a \to 0$:



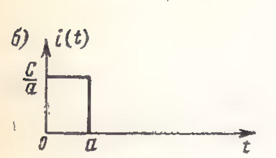


Рис. 1.18. Приближенное представление единичного скачка напряжения:
а) неидеальный скачок; *б*) импульс тока в конденсаторе, вызванный наклонным участком неидеального скачка

$$\delta(t) = \lim_{a \to 0} \frac{d}{dt}[u_a(t)] =$$

$$= \lim_{a \to 0} \frac{1}{a}\,[u_1(t) - u_1(t-a)].$$

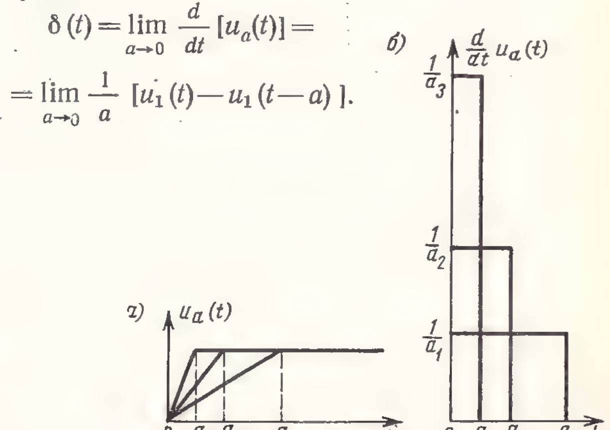


Рис. 1.19. Функция и ее производная:
а) неидеальный единичный скачок; *б*) производные, соответствующие различным наклонам неидеального скачка

Функция $\delta(t)$ равна нулю везде, за исключением точки $t=0$. Это подтверждается определением единичного импульса, данным Дираком:

$$\left.\begin{aligned} &\int_{-\infty}^{\infty} \delta(t)\,dt = 1 \\ &\delta(t) = 0 \ \text{при} \ t \neq 0 \end{aligned}\right\}. \qquad (1.83)$$

Как следует из определения (1.83), единичный импульс не является функцией в обычном математическом смысле. В теории обобщенных функций[1]) указанный импульс определяется как последовательность регулярных функций и все операции над ним рассматриваются как операции над этой последовательностью. Единичный импульс можно определить как предел последовательностей, показанных на рис. 1.20. Все эти последовательности удовлетворяют определению (1.83).

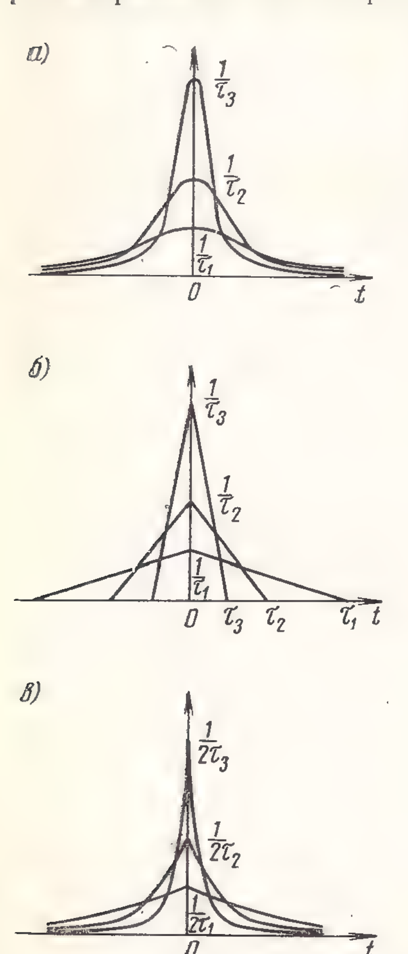


Рис. 1.20. Единичный импульс как предел последовательности: *а)* гауссовых сигналов; *б)* треугольных импульсов; *в)* двусторонних экспоненциальных сигналов

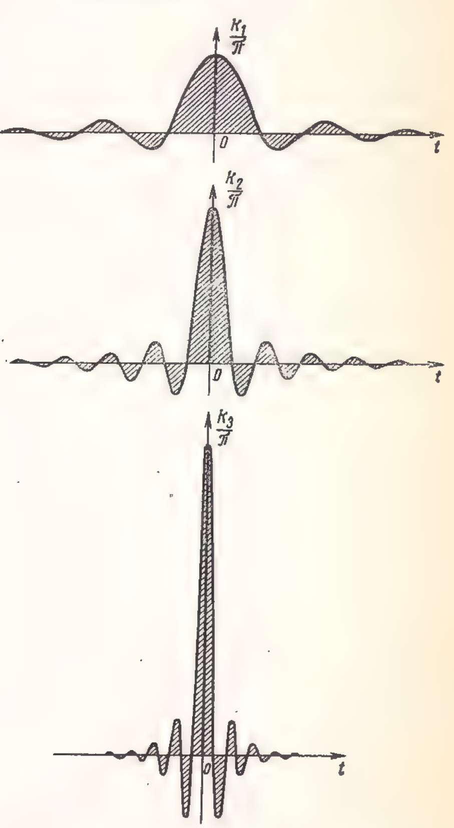


Рис. 1.21. Единичный импульс как предел последовательности функций отсчетов

---

[1]) M. T. L i g h t h i l l. Fourier analysis and generalized function, Cambridge University Press, 1959. См. также И. М. Г е л ь ф а н д, Г. Е. Ш и л о в. Обобщенные функции и действия над ними. Физматгиз, 1958. — *Прим. ред.*

Гауссов импульс $\delta(t)=\lim\limits_{\tau\to 0}\frac{1}{\tau}e^{-\pi t^2/\tau^2}$.

Треугольный импульс $\delta(t)=\begin{cases}\lim\limits_{\tau\to 0}\frac{1}{\tau}\left[1-\frac{|t|}{\tau}\right] & \text{при } |t|<\tau,\\ 0 & \text{при } |t|>\tau.\end{cases}$

Экспоненциальный импульс $\delta(t)=\lim\limits_{\tau\to 0}\frac{1}{2\tau}e^{-|t|/\tau}$.

Функция отсчетов. Можно показать, что

$$\int\limits_{-\infty}^{\infty}\frac{k}{\pi}Sa(kt)\,dt=1. \qquad (1.84)$$

С увеличением k возрастает амплитуда $\frac{k}{\pi}Sa(kt)$, функция осциллирует быстрее и очень быстро уменьшается при удалении от начала координат (рис. 1.21). В пределе при $k\to\infty$ функция существует только в начале координат, но площадь под кривой остается равной единице [см. ф-лу (1.84)]. Следовательно,

$$\delta(t)=\lim_{k\to\infty}\frac{k}{\pi}Sa(kt). \qquad (1.85)$$

Квадратичная функция отсчетов

$$\delta(t)=\lim_{k\to\infty}\frac{k}{\pi}Sa^2(kt). \qquad (1.86)$$

Это следует из того, что при предельном переходе функция $\frac{k}{\pi}Sa^2(kt)$ ведет себя так же, как $\frac{k}{\pi}Sa(kt)$, и, кроме того,

$$\int\limits_{-\infty}^{\infty}\frac{k}{\pi}Sa^2(kt)\,dt=1. \qquad (1.87)$$

Возвращаясь к единичному импульсу $\delta(t)$, видим, что площадь его концентрируется возле начала координат $t=0$. Следовательно, можно записать

$$\int\limits_{-\infty}^{\infty}\delta(t)dt=\int\limits_{0-}^{0+}\delta(t)\,dt=1, \qquad (1.88)$$

где 0^+ и 0^- означают произвольные малые величины t, приближающиеся к нулю справа и слева от начала координат соответственно. Так как $\delta(t)=0$ везде, кроме точки $t=0$,

$$\int\limits_{-\infty}^{\infty}f(t)\,\delta(t)\,dt=f(0)\int\limits_{-\infty}^{\infty}\delta(t)\,dt=f(0). \qquad (1.89a)$$

Можно также записать, что

$$\int_{-\infty}^{\infty} f(t)\,\delta(t-t_0)\,dt = f(t_0). \qquad (1.89б)$$

Выражения (1.89) описывают так называемое фильтрующее (или сдвигающее) свойство единичного импульса [1]).

Единичный скачок $u_1(t)$, единичный импульс $\delta(t)$ и их высшие производные известны как сингулярные (особые) функции. Этот термин применяется к функциям, являющимся разрывными или имеющим разрывные производные. Такие функции могут иметь непрерывные производные только конечного порядка. Например, параболическая функция $f(t)=at^2u_1(t)$ имеет непрерывную производную только первого порядка $[f'(t)=2atu_1(t)]$; вторая производная терпит разрыв. Следовательно, это сингулярная функция. Фактически любая функция, представленная полиномом по степеням t, является сингулярной.

1.12. Преобразования Фурье, включающие единичный импульс

ПРЕОБРАЗОВАНИЕ ФУРЬЕ ЕДИНИЧНОГО ИМПУЛЬСА

Преобразование Фурье единичного импульса имеет вид

$$\mathcal{F}[\delta(t)] = \int_{-\infty}^{\infty} \delta(t)\,e^{-i\omega t}\,dt. \qquad (1.90)$$

Из фильтрующего свойства единичного импульса (1.89а) следует, что интеграл в правой части ф-лы (1.90) равен единице, т. е.

$$\mathcal{F}[\delta(t)] = 1. \qquad (1.91)$$

Таким образом, преобразование Фурье единичного импульса равно единице. Это значит, что единичный импульс имеет равномерную спектральную плотность во всей бесконечной области частот (табл. 1.1, поз. *9*). Иными словами, единичный импульс содержит составляющие всех возможных частот с одинаковыми относительными амплитудами.

---

[1]) Здесь показано, что ф-лы (1.89) следуют из определения единичного импульса (1.83). На самом деле при строгом подходе единичный импульс определяется соотношениями (1.89). Этим способом единичный импульс определяется с помощью его интегральных свойств, а не как функция времени, определенная для каждого значения t. Можно показать, что ф-ла (1.83) не определяет однозначную функцию. См., например, A. Papoulis. Fourier integral and its applications, Mc. Graw-Hill, New York, 1962.

ПРЕОБРАЗОВАНИЕ ФУРЬЕ ПОСТОЯННОЙ ФУНКЦИИ

Найдем преобразование Фурье от функции вида $f(t)=A$. Эта функция не удовлетворяет условию абсолютной интегрируемости. Тем не менее она имеет в пределе преобразование Фурье. Рассмотрим преобразование Фурье стробирующей функции с амплитудой A и длительностью τ. В пределе при $\tau\to\infty$ стробирующая функция преобразуется в постоянную A. Поэтому можно рассматривать преобразование Фурье стробирующей функции $G_\tau(t)$ при $\tau\to\infty$. Выше было получено преобразование Фурье $AG_\tau(t)$ в виде $A\tau Sa(\omega\tau/2)$. Следовательно,

$$\mathcal{F}[A]=\lim_{\tau\to\infty} A\tau\ Sa\frac{\omega\tau}{2}=2\pi A\lim_{\tau\to\infty}\frac{\tau}{2\pi}\ Sa\left(\frac{\omega\tau}{2}\right).$$

Из ф-лы (1.85) вытекает, что предел функции отсчетов есть единичный импульс $\delta(\omega)$. Следовательно,

$$\mathcal{F}[A]=2\pi A\,\delta(\omega);\quad \mathcal{F}[1]=2\pi\delta(\omega). \tag{1.92}$$

Таким образом, если $f(t)$ равна некоторой постоянной, она содержит только одну частотную составляющую на $\omega=0$ (табл. 1.1, поз. *10*). К этому же выводу можно прийти логически, так как постоянная функция представляет собой сигнал постоянного тока ($\omega=0$), который не содержит никаких частотных составляющих, кроме составляющей на нулевой частоте.

ПРЕОБРАЗОВАНИЕ ФУРЬЕ ФУНКЦИИ sign (t)

Сигнум-функция определяется как

$$\operatorname{sign}(t)=\begin{cases}1 & \text{при } t>0,\\ -1 & \text{при } t<0.\end{cases} \tag{1.93}$$

Легко видеть, что

$$\operatorname{sign}(t)=2u_1(t)-1. \tag{1.94}$$

Преобразование Фурье $\operatorname{sign}(t)$ можно получить, записав эту функцию в виде

$$\operatorname{sign}(t)=\lim_{a\to0}\left[e^{-at}u_1(t)-e^{at}u_1(-t)\right].$$

Следовательно,

$$\mathcal{F}[\operatorname{sign}(t)]=\lim_{a\to0}\left[\int_0^\infty e^{-at}e^{-i\omega t}dt-\int_{-\infty}^0 e^{at}e^{-i\omega t}dt\right]=$$

$$=\lim_{a\to0}[-2/i\,\omega(a^2+\omega^2)]=2/(i\omega). \tag{1.95}$$

График амплитудного спектра сигнум-функции приведен в табл. 1.1 (поз. *18*).

Из ф-лы (1.94) следует, что

$$u_1(t) = \frac{1}{2}[1 + \text{sign}(t)].$$

Следовательно,

$$\mathcal{F}[u_1(t)] = \frac{1}{2}\{\mathcal{F}[1] + \mathcal{F}[\text{sign}(t)]\}.$$

Используя ф-лы (1.92) и (1.95), получаем

$$\mathcal{F}[u_1(t)] = \pi\delta(\omega) + \frac{1}{i\omega}. \tag{1.96}$$

Спектральная плотность содержит единичный импульс на частоте $\omega=0$ (табл. 1.1, поз. *15*). Как и следовало ожидать, функция $u_1(t)$ имеет постоянную составляющую и другие частотные составляющие. Функция $u_1(t)$ на первый взгляд представляется сигналом постоянного тока, и поэтому наличие других частотных составляющих, кроме составляющей с $\omega=0$, может показаться несколько странным. Однако функция $u_1(t)$ не является чистым сигналом постоянного тока. Она равна нулю при $t<0$ и в точке $t=0$ имеет скачок, порождающий другие частотные составляющие. Для того чтобы $u_1(t)$ была чистым сигналом постоянного тока, она должна быть постоянной на бесконечном интервале $(-\infty, \infty)$. Как было показано выше, [ф-ла (1.92)], такой сигнал действительно не имеет частотных составляющих, кроме составляющей на частоте $\omega=0$.

БЕСКОНЕЧНЫЕ ГАРМОНИЧЕСКИЕ СИГНАЛЫ $\cos\omega_0 t$ и $\sin\omega_0 t$

Рассмотрим гармонические сигналы $\cos\omega_0 t$ и $\sin\omega_0 t$, существующие на бесконечном интервале $(-\infty, \infty)$. Эти сигналы не удовлетворяют условию абсолютной интегрируемости, однако их преобразования Фурье существуют и их можно найти, если использовать прием временного ограничения процесса, примененный к постоянной функции $f(t)=A$. Предположим сначала, что эти функции существуют на интервале $(-\tau/2, \tau/2)$, а вне этого интервала равны нулю. Затем в пределе устремим τ к бесконечности. Продемонстрируем теперь этот прием:

$$\mathcal{F}(\cos\omega_0 t) = \lim_{\tau\to\infty}\int_{-\tau/2}^{\tau/2}\cos\omega_0 t\, e^{-i\omega t}\,dt =$$

$$= \lim_{\tau\to\infty}\frac{\tau}{2}\left\{\frac{\sin[(\omega-\omega_0)\tau/2]}{(\omega-\omega_0)\tau/2} + \frac{\sin[(\omega+\omega_0)\tau/2]}{(\omega+\omega_0)\tau/2}\right\} =$$

$$= \lim_{\tau\to\infty}\left\{\frac{\tau}{2} Sa\left[\frac{\tau(\omega-\omega_0)}{2}\right] + \frac{\tau}{2} Sa\left[\frac{\tau(\omega+\omega_0)}{2}\right]\right\}. \tag{1.97}$$

В пределе функция отсчетов согласно ф-ле (1.85) переходит в единичный импульс:

$$\mathcal{F}(\cos\omega_0 t) = \pi[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]. \tag{1.98}$$

Аналогично можно получить

$$\mathcal{F}(\sin\omega_0 t) = \mathrm{i}\,\pi[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)] \tag{1.99}$$

Спектры этих функций состоят из двух единичных импульсов на частотах ω_0 и $-\omega_0$ (табл. 1.1, поз. *11, 12*). Интересно проследить за изменением спектра при постепенном увеличении τ до бесконечности. Для конечного τ спектральная плотность определяется выражением (1.97). График ее для $\tau=16\pi/\omega_0$ приведен на рис. 1.22. Этот график спектральной плотности соответствует сигналу $\cos\omega_0 t$, имеющему только восемь периодов:

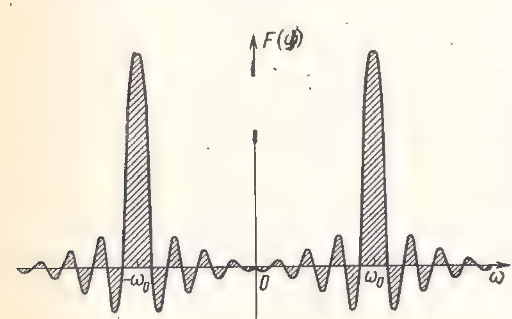


Рис. 1.22. Спектральная плотность отрезка сигнала $\cos\omega_0 t$, состоящего из восьми периодов

$$f(t) = \begin{cases} \cos\omega_0 t \ \text{при} \ |t| < \dfrac{16\pi}{\omega_0}, \\ 0 \quad \text{при} \ |t| > \dfrac{16\pi}{\omega_0}. \end{cases}$$

Заметим, что основная энергия сосредоточена на частотах вблизи $\pm\omega_0$. В пределе при $\tau\to\infty$ спектральная плотность равна нулю везде, за исключением частот $\pm\omega_0$, где она бесконечна, но площадь под кривой на этих частотах равна π. Поэтому в пределе спектр представляет собой два единичных импульса с площадью в π единиц каждый, расположенных на частотах $\pm\omega_0$, как это показано на рисунке поз. *11* табл. 1.1. Очевидно, спектральная плотность функций $\cos\omega_0 t$ и $\sin\omega_0 t$ существует только на частоте $\omega=\omega_0$, так как эти сигналы не содержат других частотных составляющих. С другой стороны, функции $\cos\omega_0 t\, u_1(t)$ и $\sin\omega_0 t\, u_1(t)$ содержат составляющие с частотами, отличными от ω_0. Можно показать, что

$$\mathcal{F}[(\cos\omega_0 t)\,u_1(t)] = \frac{\pi}{2}[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)] + \frac{\mathrm{i}\,\omega}{\omega_0^2-\omega^2} \cdot \tag{1.100a}$$

$$\mathcal{F}[(\sin\omega_0 t)\,u_1(t)] = \frac{\pi}{2\mathrm{i}}[\delta(\omega-\omega_0)-\delta(\omega+\omega_0)] + \frac{\omega_0}{\omega_0^2-\omega^2}. \tag{1.100б}$$

Эти функции, очевидно, содержат составляющую частоты ω_0, а также составляющие других частот (табл. 1.1, поз. *13, 14*).

На первый взгляд представляется, что $\cos\omega_0 t\, u_1(t)$ и $\sin\omega_0 t\, u_1(t)$ — чисто гармонические сигналы, и может показаться странным, что они содержат составляющие с частотами, отличными от ω_0. Необходимо помнить, однако, что функция здесь выражается с помощью бесконечных экспоненциальных (или бесконечных синусоидальных) функций на интервале $(-\infty, \infty)$. Функции же $\cos\omega_0 t\, u_1(t)$ и $\sin\omega_0 t\, u_1(t)$ не являются бесконечными синусоидальными сигналами. Они равны нулю при $t<0$ и существуют только при положительных значениях t, а следовательно, кроме ω_0, содержат также и другие частотные составляющие. Бесконечное число частотных составляющих в спектре этих функций суммируется таким образом, что при $t<0$ функция равна нулю, а при $t>0$ равна $\cos\omega_0 t$ (или $\sin\omega_0 t$).

Если же синусоидальные сигналы существуют на бесконечном интервале, то, как показывают ф-лы (1.98) и (1.99), они действительно имеют составляющие только на частотах $\pm\omega_0$.

ПРЕОБРАЗОВАНИЕ ФУРЬЕ БЕСКОНЕЧНОЙ ЭКСПОНЕНТЫ $e^{i\omega_0 t}$

Найдем преобразование Фурье бесконечного экспоненциального сигнала $e^{i\omega_0 t}$ $(-\infty<t<\infty)$. Имеем $e^{i\omega_0 t}=\cos\omega_0 t+i\sin\omega_0 t$. Следовательно, $\mathcal{F}[e^{i\omega_0 t}]=\mathcal{F}[\cos\omega_0 t+i\sin\omega_0 t]$.

Подставив в это соотношение ф-лы (1.98) и (1.99), получим

$$\mathcal{F}[e^{i\omega_0 t}]=\pi[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)-\delta(\omega+\omega_0)+\delta(\omega-\omega_0)]=2\pi\delta(\omega-\omega_0) \qquad (1.101)$$

Таким образом, преобразование Фурье сигнала $e^{i\omega_0 t}$ есть умноженный на 2π единичный импульс на частоте ω_0. Отметим, что сигнал $e^{i\omega_0 t}$ не является действительной функцией времени и поэтому имеет спектр в виде только одной составляющей на частоте $\omega=\omega_0$. Выше было показано, что для любой действительной функции времени спектральная плотность $F(\omega)$ удовлетворяет условию [см. ф-лу (1.77)]: $F^*(\omega)=F(-\omega)$, $|F(\omega)|=|F(-\omega)|$. Следовательно, амплитудный спектр любой действительной функции времени есть четная функция частоты, и если имеется единичный импульс на $\omega=\omega_0$, то должен существовать единичный импульс и на частоте $\omega=-\omega_0$. Примерами этого служат спектры сигналов $\sin\omega_0 t$ и $\cos\omega_0 t$.

ПРЕОБРАЗОВАНИЕ ФУРЬЕ ПЕРИОДИЧЕСКОЙ ФУНКЦИИ

Выше было получено преобразование Фурье как предельный случай ряда Фурье, когда период периодической функции полагается бесконечным. Теперь займемся противоположным анализом и покажем, что ряд Фурье является предельным случаем преобразования Фурье. Такое представление периодической функции полезно, так как обеспечивает единый аппарат анализа как периодических, так и непериодических функций.

Строго говоря, преобразование Фурье для периодической функции не существует, так как последняя не удовлетворяет условию абсолютной интегрируемости. Для любой периодической функции $f(t)$.

$$\int_{-\infty}^{\infty} |f(t)|\, dt = \infty.$$

Но в пределе преобразование существует (например, преобразование Фурье сигналов $\cos\omega_0 t$ и $\sin\omega_0 t$). Воспользуемся и теперь этим же приемом, предположив, что периодическая функция существует только на конечном интервале $(-\tau/2, \tau/2)$, а затем устремив τ к бесконечности.

С другой стороны, периодическую функцию можно представить рядом Фурье. Тогда преобразование Фурье периодической функции равно сумме преобразований Фурье отдельных составляющих. Запишем периодическую функцию $f(t)$ с периодом T в виде:

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{in\omega_0 t}, \quad \omega_0 = 2\pi/T.$$

Возьмем преобразование Фурье от обеих частей равенства:

$$\mathcal{F}[f(t)] = \mathcal{F} \sum_{n=-\infty}^{\infty} F_n e^{in\omega_0 t} = \sum_{n=-\infty}^{\infty} F_n \mathcal{F}\left(e^{in\omega_0 t}\right).$$

Подставив преобразование $e^{i\omega_0 t}$ из ф-лы (1.101), получим

$$\mathcal{F}[f(t)] = 2\pi \sum_{n=-\infty}^{\infty} F_n \delta(\omega - n\omega_0). \qquad (1.102)$$

Выражение (1.102) устанавливает, что *функция спектральной плотности или преобразование Фурье периодической функции состоит из единичных импульсов, расположенных на частотах гармоник сигнала, с интенсивностями в 2π раз больше соответствующих коэффициентов экспоненциального ряда Фурье.* Последовательность равноотстоящих единичных импульсов как раз и есть предельная форма непрерывной функции спектральной плотности. Этот вывод, конечно, не должен быть неожиданным, поскольку уже известно, что периодическая функция содержит только составляющие с частотами гармоник.

Пример 1.7

Найдем преобразование Фурье периодической стробирующей функции (т. е. прямоугольного импульса длительностью τ *сек*, повторяющегося каждые T сек). Ряд Фурье для этой функции был получен ранее [ф-ла (1.66*б*)]:

$$f(t) = \sum_{n=-\infty}^{\infty} F_n e^{in\omega_0 t},$$

где

$$F_n = \frac{A\tau}{T} Sa\left(\frac{n\pi\tau}{T}\right).$$

Из ф-лы (1.102) следует, что преобразование Фурье этой функции равно

$$\mathscr{F}[f(t)] = \frac{2\pi A\tau}{T} \sum_{n=-\infty}^{\infty} Sa\left(\frac{n\pi\tau}{T}\right)\delta(\omega - n\omega_0). \qquad (1.103)$$

Преобразование Фурье рассматриваемой функции состоит из единичных импульсов, расположенных на частотах $\omega=0$, $\pm\omega_0$, $\pm 2\omega_0$, ..., $\pm n\omega_0$ ... и т. д.

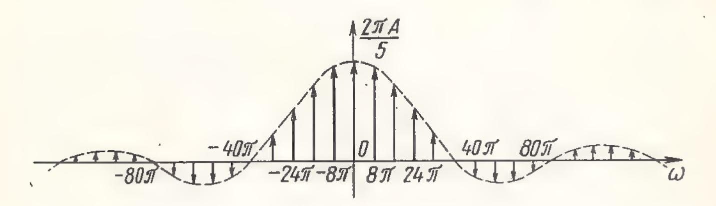


Рис. 1.23. Спектральное представление периодической стробирующей функции

Интенсивность (площадь) единичного импульса на частоте $\omega=n\omega_0$ равна $2\pi(A\tau/T)Sa(\pi n\tau/T)$. На рис. 1.23 приведен график спектра для конкретного случая $\tau=1/20$ *сек.* Здесь $\omega_0=8\pi$.

Пример 1.8

Найдем преобразование Фурье периодической последовательности единичных импульсов, повторяющихся через T *сек* (табл. 1.1, поз. 16). Эта функция занимает важное место в теории временной дискретизации, и для нее следует ввести специальное обозначение $\delta_T(t)$. Таким образом,

$$\delta_T(t) = \delta(t) + \delta(t-T) + \delta(t-2T) + \ldots + \delta(t-nT) + \ldots +$$

$$+ \delta(t+T) + \delta(t+2T) + \ldots + \delta(t+nT) + \ldots = \sum_{n=-\infty}^{\infty} \delta(t-nT). \qquad (1.104)$$

Ряд Фурье для этой периодической функции записывается в виде

$$\delta_T(t) = \sum_{n=-\infty}^{\infty} F_n e^{i n\omega_0 t},$$

где

$$F_n = \frac{1}{T}\int_{-T/2}^{T/2} \delta_T(t)\, e^{-i n\omega_0 t}\, dt.$$

На интервале $(-T/2, T/2)$ функция $\delta_T(t)$ есть единичный импульс $\delta(t)$. Следовательно,

$$F_n = \frac{1}{T}\int_{-T/2}^{T/2} \delta(t)\, e^{-i n\omega_0 t}\, dt.$$

В силу фильтрующего свойства единичного импульса [ф-ла (1.89)] полученное соотношение оказывается равным $F_n=1/T$, т. е. постоянной величине. Это означает, что периодическая последовательность единичных импульсов с

периодом T содержит составляющие с частотами $\omega=0, \pm\omega_0, \pm2\omega_0, \ldots, \pm n\omega_0 \ldots$ и т. д. ($\omega_0=2\pi/T$) одинаковой амплитуды

$$\delta_T(t)=\frac{1}{T}\sum_{n=-\infty}^{\infty} e^{i n \omega_0 t}. \qquad (1.105)$$

Для того чтобы найти преобразование Фурье от $\delta_T(t)$, воспользуемся соотношением (1.102). Так как в данном случае $F_n=1/T$,

$$\mathscr{F}[\delta_T(t)]=2\pi\sum_{n=-\infty}^{\infty}\frac{1}{T}\delta(\omega-n\omega_0)=\frac{2\pi}{T}\sum_{n=-\infty}^{\infty}\delta(\omega-n\omega_0)=$$

$$=\omega_0\sum_{n=-\infty}^{\infty}\delta(\omega-n\omega_0)=\omega_0\,\delta_{\omega_0}(\omega). \qquad (1.106)$$

Соотношение (1.106) устанавливает, что преобразование Фурье периодической последовательности единичных импульсов с периодом T есть последовательность единичных импульсов с одинаковыми амплитудами, разделенных интервалами ω_0 *рад* ($\omega_0=2\pi/T$). На рис. 1.24 приведены последовательности единичных импульсов с периодами $T=1/2$ и 1 *сек* и соответствующие преобразования Фурье. Очевидно, при увеличении периода частотный спектр становится плотнее.

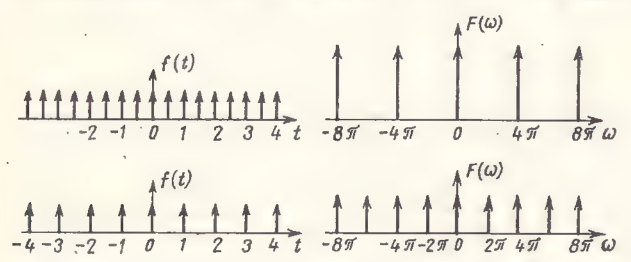


Рис. 1.24. Периодические последовательности единичных функций и их преобразования Фурье

1.13. Некоторые свойства преобразования Фурье

ВВОДНЫЕ ЗАМЕЧАНИЯ

Преобразование Фурье служит инструментом для представления функции времени через экспоненциальные составляющие с различными частотами. Ранее уже указывалось, что преобразование Фурье является одним из способов определения функции. Итак, известны два способа описания функции — во временно́й и частотной областях. Выясним, как влияют определенные действия, совершаемые над функцией в одной области, на ее представление в другой. Например, как связаны между собой спектры функции и ее производной, если функция дифференцируется во временно́й области, или что произойдет со спектром функции, если функцию сдвинуть во времени?

В уравнениях, определяющих функцию в обеих областях, существует определенная симметрия. В этом легко убедиться, рассмотрев пару преобразований Фурье [ф-лы (1.73), (1.74)]. Поэтому следует ожидать, что указанная симметрия должна отражаться в свойствах преобразований. Например, эффект в частотной области обусловленный дифференцированием функции во временной области, должен быть подобен эффекту во временно́й области, обусловленному дифференцированием в частотной области.

Убедимся, что это действительно так. Для удобства обозначим соответствие между двумя областями двойной стрелкой:

$$f(t) \longleftrightarrow F(\omega), \tag{1.107}$$

Последнее обозначает, что $F(\omega)$ есть прямое преобразование Фурье $f(t)$, а $f(t)$ есть обратное преобразование Фурье $F(\omega)$ в соответствии с ф-лами (1.73) и (1.74).

СВОЙСТВО СИММЕТРИИ

Если $f(t) \longleftrightarrow F(\omega)$, то

$$F(t) \longleftrightarrow 2\pi f(-\omega). \tag{1.108}$$

Для доказательства этого свойства воспользуемся ф-лами (1.73) и (1.74), из которых следует, что

$$2\pi f(-t) = \int\limits_{-\infty}^{\infty} F(\omega)\, e^{-i\omega t}\, d\omega.$$

Так как ω в этом интеграле является произвольной переменной, ее можно заменить на x. Поэтому

$$2\pi f(-t) = \int\limits_{-\infty}^{\infty} F(x)\, e^{-ixt}\, dx.$$

Следовательно,

$$2\pi f(-\omega) = \int\limits_{-\infty}^{\infty} F(x)\, e^{-ix\omega}\, dx.$$

Заменив теперь произвольную переменную x другой переменной t, получим

$$2\pi f(-\omega) = \int\limits_{-\infty}^{\infty} F(t)\, e^{-i\omega t}\, dt = \mathcal{F}\,[F(t)].$$

Следовательно, $F(t) \longleftrightarrow 2\pi f(-\omega)$. Если $f(t)$ — четная функция, то $f(-\omega) = f(\omega)$ и

$$F(t) \longleftrightarrow 2\pi f(\omega). \tag{1.109}$$

Это свойство иллюстрируется рис. 1.25, из которого видно, что преобразование Фурье стробирующей функции представляет собой функцию отсчетов, а преобразование Фурье функции отсчетов есть

стробирующая функция. Свойство симметрии справедливо для всех (а не только для четных) функций. Если $f(t)$ — нечетная функция, симметрия не столь наглядна, но тем не менее она существует, как это следует из ф-лы (1.108).

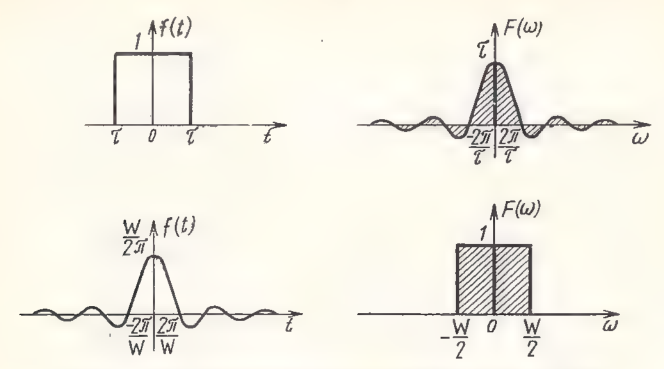


Рис. 1.25. Симметричность преобразования Фурье

СВОЙСТВО ЛИНЕЙНОСТИ

Если $f_1(t) \longleftrightarrow F_1(\omega)$ и $f_2(t) \longleftrightarrow F_2(\omega)$, то для любых произвольных постоянных a_1 и a_2

$$a_1 f_1(t) + a_2 f_2(t) \longleftrightarrow a_1 F_1(\omega) + a_2 F_2(\omega). \tag{1.110}$$

Доказательство тривиально. Свойство линейности сохраняется для конечной суммы:

$$a_1 f_1(t) + a_2 f_2(t) + \ldots + a_n f_n(t) \longleftrightarrow a_1 F_1(\omega) + a_2 F_2(\omega) + \ldots + a_n F_n(\omega)$$

СВОЙСТВО ИЗМЕНЕНИЯ МАСШТАБА

Если $f(t) \longleftrightarrow F(\omega)$, то для любой действительной постоянной a

$$f(at) \longleftrightarrow \frac{1}{|a|} F\left(\frac{\omega}{a}\right). \tag{1.111}$$

Докажем это для действительной положительной постоянной a:

$$\mathcal{F}[f(at)] = \int_{-\infty}^{\infty} f(at) e^{-i\omega t} dt.$$

Пусть $x = at$. Тогда

$$\mathcal{F}[f(at)] = \frac{1}{a} \int_{-\infty}^{\infty} f(x) e^{-i\omega x/a} dx = \frac{1}{a} F\left(\frac{\omega}{a}\right).$$

Следовательно, $f(at) \longleftrightarrow \frac{1}{a} F\left(\frac{\omega}{a}\right)$.

Подобным образом можно показать, что для $a<0$

$$f(at) \longleftrightarrow \frac{1}{-a} F\left(\frac{\omega}{a}\right).$$

Следовательно, $f(at) \longleftrightarrow \frac{1}{|a|} F\left(\frac{\omega}{a}\right)$.

Функция $f(at)$ представляет собой функцию $f(t)$, сжатую по временной шкале в a раз[1]). Точно так же функция $F(\omega/a)$ представляет собой функцию $F(\omega)$, растянутую по шкале частот в a раз. Сжатая во времени функция изменяется в a раз быстрее и, следовательно, частоты ее составляющих должны в a раз увеличиться.

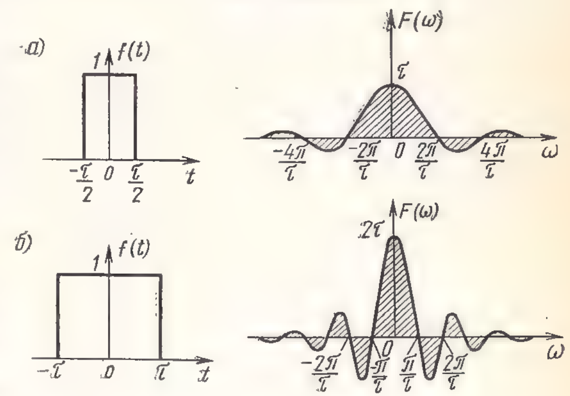


Рис. 1.26. К пояснению сжатия функции во времени:
а) стробирующая функция длительностью T и ее преобразование Фурье;
б) стробирующая функция длительностью $2T$ и ее преобразование Фурье

В качестве примера рассмотрим сигнал $\cos\omega_0 t$. Этот сигнал имеет частотные составляющие на $\pm\omega_0$. Сигнал $\cos 2\omega_0 t$ представляет собой сжатый вдвое сигнал $\cos\omega_0 t$ и его частотные составляющие располагаются на $\pm 2\omega_0$.

Свойство изменения масштаба иллюстрируется рис. 1.26. В особом случае, когда $a=-1$, согласно свойству изменения масштаба имеем

$$f(-t) \longleftrightarrow F(-\omega). \tag{1.112}$$

Пример 1.9

Найдем обратное преобразование Фурье функций $\text{sign}(\omega)$ и $u_1(\omega)$, воспользовавшись для этого свойствами симметрии и изменения масштаба.

Из ф-лы (1.95) имеем $\text{sign}(t) \longleftrightarrow 2/i\omega$. Применив к этому соотношению свойство симметрии (1.108), получим

$$2/it \longleftrightarrow 2\pi\,\text{sign}(-\omega).$$

Но $\text{sign}(-\omega) = -\text{sign}(\omega)$, следовательно,

$$i/\pi t \longleftrightarrow \text{sign}(\omega). \tag{1.113}$$

---

[1]) Предполагается, что $a>1$; если $a<1$, то функция «растягивается» во времени. — *Прим. ред.*

Единичный скачок $u_1(\omega)$ можно записать в виде $u_1(\omega) = {}^1/_2[1 + \text{sign}(\omega)]$. Следовательно,

$$\mathscr{F}^{-1}[u_1(\omega)] = \frac{1}{2}\{\mathscr{F}^{-1}[1] + \mathscr{F}^{-1}[\text{sign}(\omega)]\} = \frac{1}{2}\delta(t) + \frac{\text{i}}{2\pi t}.$$

Таким образом,

$$\left[\frac{1}{2}\delta(t) - \frac{1}{2\pi \text{i} t}\right] \longleftrightarrow u_1(\omega). \qquad (1.114)$$

СВОЙСТВО ЧАСТОТНОГО СДВИГА

Если $f(t) \longleftrightarrow F(\omega)$, то

$$f(t)\,\text{e}^{\text{i}\omega_0 t} \longleftrightarrow F(\omega - \omega_0). \qquad (1.115)$$

Действительно,

$$\mathscr{F}\left[f(t)\,\text{e}^{\text{i}\omega_0 t}\right] = \int_{-\infty}^{\infty} f(t)\,\text{e}^{\text{i}\omega_0 t}\,\text{e}^{-\text{i}\omega t}\,dt = \int_{-\infty}^{\infty} f(t)\,\text{e}^{-\text{i}(\omega - \omega_0)t}\,dt = F(\omega - \omega_0).$$

Таким образом, сдвиг на ω_0 в частотной области эквивалентен умножению на $\text{e}^{\text{i}\omega_0 t}$ во временно́й области. Очевидно, умножение на $\text{e}^{\text{i}\omega_0 t}$ переносит весь спектр $F(\omega)$ на частоту ω_0. Поэтому данная теорема известна также под названием *теоремы о смещении спектра.*

Необходимость переноса частотного спектра часто возникает в системах связи. Обычно перенос спектра осуществляется перемножением сигнала $f(t)$ на синусоидальный сигнал. Этот процесс известен как *модуляция*[1]. Поскольку синусоидальный сигнал частоты ω_0 можно выразить суммой экспонент, очевидно, перемножение $f(t)$ на синусоидальный сигнал будет смещать весь частотный спектр сигнала. Последнее легко показать с помощью тождества

$$f(t)\cos\omega_0 t = \frac{1}{2}\left[f(t)\,\text{e}^{\text{i}\omega_0 t} + f(t)\,\text{e}^{-\text{i}\omega_0 t}\right].$$

Если $f(t) \longleftrightarrow F(\omega)$, то на основании теоремы о смещении спектра

$$f(t)\cos\omega_0 t \longleftrightarrow \frac{1}{2}[F(\omega + \omega_0) + F(\omega - \omega_0)]. \qquad (1.116а)$$

Подобным образом можно показать, что

$$f(t)\sin\omega_0 t \longleftrightarrow \frac{\text{i}}{2}[F(\omega + \omega_0) - F(\omega - \omega_0)]. \qquad (1.116б)$$

---

[1]) Сдвиг спектра называется также преобразованием частоты сигнала. — *Прим. ред.*

Таким образом, в процессе модуляции осуществляется перенос спектра на $\pm\omega_0$. Это весьма важный результат для теории связи.

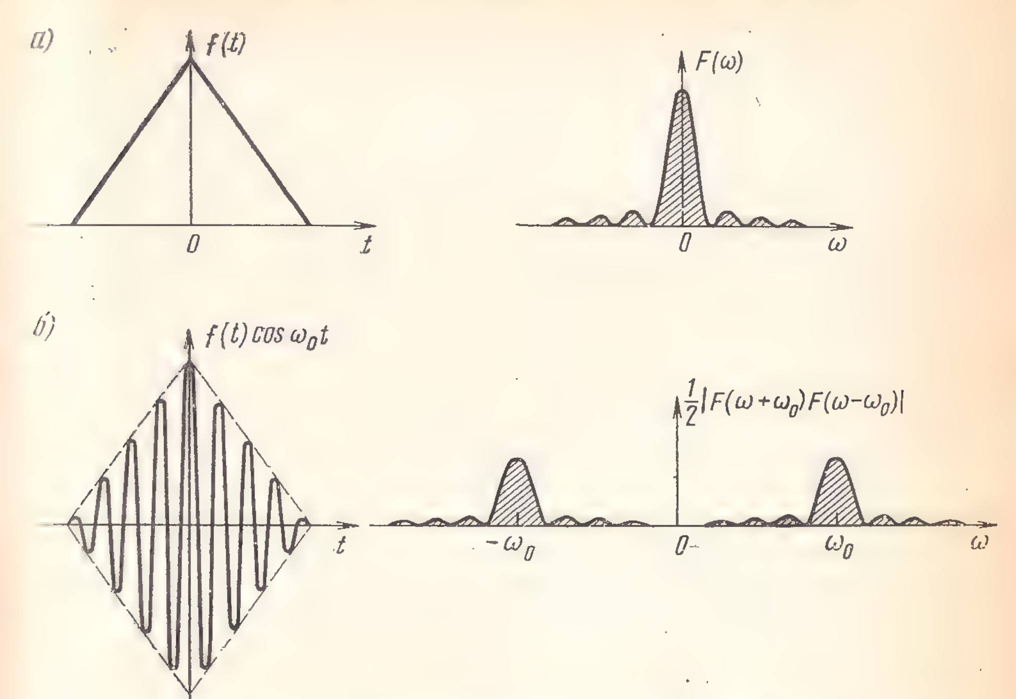


Рис. 1.27. Перенос спектра при модуляции:
а) треугольный импульс и его спектр; *б)* модулированный сигнал и его спектр

На рис. 1.27 показан пример переноса спектра вследствие модуляции. Рассмотренную теорему часто называют также *теоремой о модуляции.*

СВОЙСТВО ВРЕМЕННОГО СДВИГА

Если $f(t) \longleftrightarrow F(\omega)$, то

$$f(t-t_0) \longleftrightarrow F(\omega)\, e^{-i\omega t_0}\,. \tag{1.117}$$

Для доказательства этого свойства воспользуемся выражением

$$\mathcal{F}\,[f(t-t_0)] = \int_{-\infty}^{\infty} f(t-t_0)\, e^{-i\omega t}\, dt.$$

Пусть $t-t_0=x$. Тогда

$$\mathcal{F}\,[f(t-t_0)] = \int_{-\infty}^{\infty} f(x)\, e^{-i\omega(x+t_0)}\, dx = F(\omega)\, e^{-i\,\omega t_0}$$

Данная теорема устанавливает, что при временном сдвиге функции на t_0 ее амплитудный спектр $|F(\omega)|$ не изменяется. Изменения происходят только в фазовом спектре на величину $-\omega t_0$. Действительно, сдвиг функции во времени означает сдвиг на t_0 каждой составляющей, а не величины самого сигнала. Сдвиг во времени на t_0 для составляющей с частотой ω эквивалентен фазовому сдвигу на $-\omega t_0$.

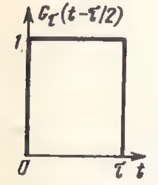


Рис. 1.28. Прямоугольный импульс

Пример 1.10

Найдем преобразование Фурье прямоугольного импульса, изображенного на рис. 1.28. Импульс представляет собой стробирующую функцию $G_\tau(t)$, сдвинутую относительно начала координат на $\tau/2$ *сек.*, поэтому ее можно записать как $G_\tau(t-\tau/2)$.

На основании табл. 1.1 и свойства временно́го сдвига имеем

$$G_\tau(t-\tau/2) \longleftrightarrow \tau\, Sa(\omega\tau/2)\, e^{-i\omega\tau/2}.$$

Обратим внимание на двойственность свойств частотного и временно́го сдвигов. Такая же двойственность имеет место и в теореме о модуляции:

$$\frac{1}{2}[f(t+T)+f(t-T)] \longleftrightarrow F(\omega)\cos\omega T. \qquad (1.118)$$

Читателю предоставляется возможность самостоятельно доказать последнее соотношение.

ДИФФЕРЕНЦИРОВАНИЕ И ИНТЕГРИРОВАНИЕ ПО ВРЕМЕНИ

Если $f(t) \longleftrightarrow F(\omega)$, то[1])

$$\frac{df}{dt} \longleftrightarrow (i\omega)F(\omega), \qquad (1.119а)$$

и

$$\int_{-\infty}^{t} f(\tau)\,d\tau \longleftrightarrow \frac{1}{i\omega}F(\omega) \qquad (1.119б)$$

при условии, что $F(\omega)/\omega$ при $\omega=0$ имеет конечное значение[2]). Это эквивалентно утверждению, что $F(0)=0$ или $\int_{-\infty}^{\infty} f(t)\,dt=0$.

Для доказательства этого свойства воспользуемся выражением (1.73) и продифференцируем его по времени:

$$\frac{df}{dt}=\frac{1}{2\pi}\frac{d}{dt}\int_{-\infty}^{\infty} F(\omega)\,e^{i\omega t}\,d\omega.$$

---

[1]) Соотношение (1.119а) не гарантирует существования преобразования Фурье df/dt. Оно просто утверждает, что если преобразование Фурье существует, то оно равно $i\omega F(\omega)$.

[2]) Если это условие не выполняется, то ф-ла (1.119б) изменяется (см. задачу 1.42).

Изменив порядок дифференцирования и интегрирования, получим

$$\frac{df}{dt}=\frac{1}{2\pi}\int_{-\infty}^{\infty} i\,\omega F(\omega)\,e^{i\omega t}\,d\omega.$$

Из этого соотношения очевидно, что $df/dt \longleftrightarrow i\omega F(\omega)$.

Аналогичным образом можно получить результат для n-й производной

$$\frac{d^n f}{dt^n} \longleftrightarrow (i\,\omega)^n F(\omega). \tag{1.120}$$

Рассмотрим теперь функцию $\varphi(t)=\int_{-\infty}^{t} f(\tau)\,d\tau$. Тогда $d\varphi(t)/dt=f(t)$.

Следовательно, если $\varphi(t)\longleftrightarrow\Phi(\omega)$, то $f(t)\longleftrightarrow i\,\omega\,\Phi(\omega)$, т. е. $F(\omega)=$ $=i\,\omega\Phi(\omega)$. Поэтому

$$\Phi(\omega)=\frac{1}{i\,\omega}F(\omega)$$

и, таким образом,

$$\int_{-\infty}^{t} f(\tau)\,d\tau \longleftrightarrow \frac{1}{i\,\omega}F(\omega).$$

Отметим, что полученный результат имеет место лишь в случае существования $\Phi(\omega)$, т. е. если $\varphi(t)$ — абсолютно интегрируемая функция. Последнее возможно только при $\lim\limits_{t\to\infty}\varphi(t)=0$ или $\int_{-\infty}^{\infty} f(t)\,dt=0$. Это эквивалентно условию $F(0)=0$, так как $\int_{-\infty}^{\infty} f(t)\,dt=F(\omega)\,|_{\omega=0}$

Теоремы о дифференцировании и интегрировании по времени (1.119а) и (1.119б) можно доказать также путем качественных рассуждений. Преобразование Фурье выражает функцию $f(t)$ в виде бесконечной суммы экспоненциальных функций вида $e^{i\omega t}$. Поэтому производная функции $f(t)$ равна непрерывной сумме производных отдельных экспоненциальных составляющих. Но производная от $e^{i\omega t}$ есть $i\,\omega\, e^{i\omega t}$, поэтому процесс дифференцирования $f(t)$ эквивалентен умножению на $i\omega$ каждой экспоненциальной составляющей. Следовательно,

$$df/dt \longleftrightarrow i\,\omega F(\omega).$$

Аналогичным образом можно рассуждать и в отношении интегрирования.

Итак, можно сделать вывод, что дифференцирование во временно́й области эквивалентно умножению в частотной области на

iω, а интегрирование во временной области эквивалентно делению в частотной области на iω. С помощью теоремы о дифференцировании во временной области можно вывести преобразование Фурье для некоторых кусочно-непрерывных функций, что иллюстрируется следующим примером.

Пример 1.11

Найдем преобразование Фурье трапецеидальной функции $f(t)$ (рис. 1.29). Дважды продифференцировав эту функцию, получим последовательность единичных импульсов, преобразование Фурье которых было найдено ранее. Из рис. 1.29 видно, что

$$\frac{d^2 f}{dt^2} = \frac{A}{b-a}\left[\delta(t+b) - \delta(t+a) - \delta(t-a) + \delta(t-b)\right]. \qquad (1.121)$$

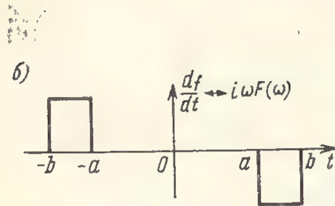
Преобразование Фурье единичного импульса равно единице, поэтому на основании теоремы о временном сдвиге имеем $\delta(t-t_0) \longleftrightarrow e^{-i\omega t_0}$. Используя этот результат и теорему о дифференцировании во временной области, преобразование (1.121) можно записать в виде

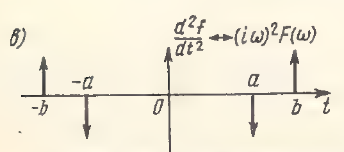
$$(i\omega)^2 F(\omega) = \frac{A}{b-a}\left(e^{i\omega b} - e^{i\omega a} - e^{-i\omega a} + e^{-i\omega b}\right),$$

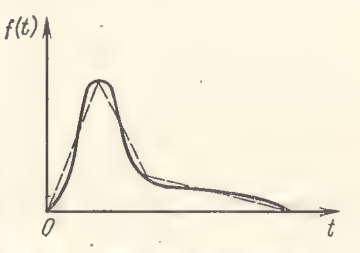
откуда следует, что

$$F(\omega) = \frac{2A}{b-a}\,\frac{\cos a\omega - \cos b\omega}{\omega^2}.$$

Эта задача подсказывает численный метод нахождения преобразования Фурье функции $f(t)$. Любую функцию, как это показано на рис. 1.30, можно аппроксимировать отрезками прямых. С увеличением числа отрезков точность аппроксимации улучшается. Вторая производная аппроксимированной функции имеет вид последовательности







в) $\frac{d^2f}{dt^2} \leftrightarrow (i\omega)^2 F(\omega)$

Рис. 1.29. Преобразование Фурье кусочно-непрерывных функций: *а*) трапецеидальная функция; *б*) ее первая производная; *в*) ее вторая производная

Рис. 1.30. Аппроксимация непрерывной функции отрезками прямых

единичных импульсов, преобразование Фурье которых легко находится. Преобразование $F(\omega)$ заданной функции равно произведению преобразования Фурье второй производной и $1/(i\omega)^2$.

Если $f(t) \longleftrightarrow F(\omega)$, то

$$-\mathrm{i}\,tf(t) \longleftrightarrow dF/d\omega. \tag{1.122}$$

Для доказательства этого свойства воспользуемся соотношением (1.74) и продифференцируем его по частоте:

$$\frac{dF}{d\omega} = \frac{d}{d\omega}\int_{-\infty}^{\infty} f(t)\,\mathrm{e}^{-\mathrm{i}\omega t}\,dt.$$

Изменив порядок дифференцирования и интегрирования, получим

$$\frac{dF}{d\omega} = \int_{-\infty}^{\infty} -\mathrm{i}\,tf(t)\,\mathrm{e}^{-\mathrm{i}\omega t}\,dt.$$

Из этого соотношения очевидно, что $-\mathrm{i}tf(t) \longleftrightarrow dF/d\omega$. Для высших производных $F(\omega)$ можно получить: $(-\mathrm{i}t)^n f(t) = d^nF/d\omega^n$.

Итак, дифференцирование в частотной области эквивалентно умножению на $-\mathrm{i}t$ во временно́й области.

ТЕОРЕМА О СВЕРТКЕ

В частотном анализе теорема о свертке является, пожалуй, одним из наиболее эффективных приемов. Она позволяет легко получать важные результаты и в последующем изложении будет часто применяться. Из двух данных функций $f_1(t)$ и $f_2(t)$ образуется интеграл

$$f(t) = \int_{-\infty}^{\infty} f_1(\tau)\,f_2(t-\tau)\,d\tau. \tag{1.123}$$

Этот интеграл и определяется как свертка функций $f_1(t)$ и $f_2(t)$. Интеграл (1.123) символически записывается в виде

$$f(t) = f_1(t) * f_2(t). \tag{1.124}$$

Прежде чем рассматривать физическую и геометрическую интерпретацию интеграла свертки, докажем теорему. Как и ранее, имеем дело с двумя теоремами: одна относится к свертке во времени, другая — к свертке по частоте.

Свертка во времени. Если $f_1(t) \longleftrightarrow F_1(\omega)$ и $f_2(t) \longleftrightarrow F_2(\omega)$, то

$$\int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau)\, d\tau \longleftrightarrow F_1(\omega) F_2(\omega), \tag{1.125а}$$

т. е.

$$f_1(t) * f_2(t) \longleftrightarrow F_1(\omega) F_2(\omega). \tag{1.125б}$$

Докажем, что это действительно так:

$$\mathscr{F}[f_1(t) * f_2(t)] = \int_{-\infty}^{\infty} e^{-i\omega t} \left[\int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau)\, d\tau \right] dt =$$

$$= \int_{-\infty}^{\infty} f_1(\tau) \left[\int_{-\infty}^{\infty} e^{-i\omega t} f_2(t-\tau)\, dt \right] d\tau.$$

На основании свойства временно́го сдвига (1.117) интеграл в квадратных скобках в правой части этого выражения равен $F_2(\omega) e^{-i\omega\tau}$. Следовательно,

$$\mathscr{F}[f_1(t) * f_2(t)] = \int_{-\infty}^{\infty} f_1(\tau) e^{-i\omega\tau} F_2(\omega)\, d\tau = F_1(\omega) F_2(\omega).$$

Свертка по частоте. Если $f_1(t) \longleftrightarrow F_1(\omega)$ и $f_2(t) \longleftrightarrow F_2(\omega)$, то

$$f_1(t) f_2(t) \longleftrightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} F_1(u) F_2(\omega-u)\, du \tag{1.126а}$$

или

$$f_1(t) f_2(t) \longleftrightarrow \frac{1}{2\pi} [F_1(\omega) * F_2(\omega)]. \tag{1.126б}$$

Эту теорему можно доказать точно так же, как и теорему о свертке во времени в силу симметрии прямого и обратного преобразований Фурье.

Итак, свертка двух функций во временно́й области эквивалентна перемножению их спектров в частотной области, а перемножение двух функций во временно́й области эквивалентно свертке их спектров в частотной области.

1.14. Основные свойства свертки

АЛГЕБРА СВЕРТКИ

Формула свертки (1.123) наводит на мысль, что свертка является особым видом перемножения. Действительно, можно написать законы алгебры свертки по аналогии с алгеброй перемножения.

1. Коммутативный закон

$$f_1(t) * f_2(t) = f_2(t) * f_1(t). \tag{1.127}$$

Это соотношение доказывается следующим образом:

$$f_1(t) * f_2(t) = \int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau)\, d\tau.$$

Изменив переменную τ на $t-x$, получим

$$f_1(t) * f_2(t) = \int_{-\infty}^{\infty} f_2(x) f_1(t-x)\, dx = f_2(t) * f_1(t).$$

2. Дистрибутивный закон

$$f_1(t) * [f_2(t) + f_3(t)] = f_1(t) * f_2(t) + f_1(t) * f_3(t). \tag{1.128}$$

Доказательство этого соотношения тривиально.

3. Ассоциативный закон

$$f_1(t) * [f_2(t) * f_3(t) = [f_1(t) * f_2(t)] * f_3(t). \tag{1.129}$$

Это соотношение вытекает из теоремы о свертке и выражения

$$F_1(\omega)\,[F_2(\omega) F_3(\omega)] = [F_1(\omega) F_2(\omega)]\, F_3(\omega).$$

ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ СВЕРТКИ

Графическое представление свертки весьма полезно как в теории связи, так и в теории цепей, особенно при анализе линейных цепей, если входной сигнал $f(t)$ и импульсная реакция цепи $h(t)$ заданы лишь в форме графиков. Для иллюстрации изложенного допустим, что $f_1(t)$ и $f_2(t)$ имеют вид прямоугольного и трапецеидального импульсов (рис. 1.31). Найдем графически свертку этих функций. По определению [ф-ла (1.123)]

$$f_1(t) * f_2(t) = \int_{-\infty}^{\infty} f_1(\tau) f_2(t-\tau)\, d\tau.$$

В этом интеграле τ является независимой переменной. На рис. 1.31*б* показаны функции $f_1(\tau)$ и $f_2(-\tau)$. Заметим, что $f_2(-\tau)$ получена поворотом $f_2(\tau)$ относительно вертикальной оси, проходящей через начало координат. График $f_2(t-\tau)$ представляет со-

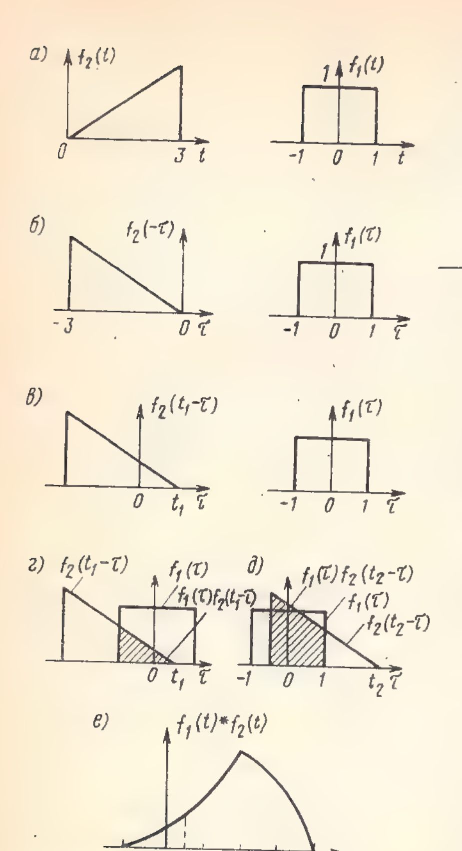


Рис. 1.31. Графическое представление свертки:
а) функции $f_2(t)$ и $f_1(t)$; *б)* зеркальные отображения функций; *в)* временной сдвиг функции $f_2(t)$ на величину t_1; *г)* определение интеграла при сдвиге t_1; *д)* определение интеграла при сдвиге t_2; *е)* график свертки функций $f_1(t)$ и $f_2(t)$

бой функцию $f_2(-\tau)$, сдвинутую на t *сек* в положительном направлении оси (рис. 1.31*в*). Величина интеграла свертки для $t=t_1$ определяется площадью под кривой произведения $f_1(\tau)f_2(t_1-\tau)$. Эта площадь заштрихована на рис. 1.31*г* и показана пунктирной линией в момент $t=t_1$ на рис. 1.31*е*. Чтобы построить график свертки как функции времени, выбираем различные значения t, сдвигаем соответственно функцию $f_2(-\tau)$ и определяем площадь под кривой произведения. В результате получается график, показанный на рис. 1.31*е*.

Заметим, что при определении свертки $f_1(t)*f_2(t)$ для положительных значений t функция $f_2(-\tau)$ сдвигается в положительном направлении оси τ, тогда как для отрицательных t функция $f_2(-\tau)$ сдвигается в отрицательном направлении оси τ.

Из ф-лы (1.127) следует, что свертка $f_1(t)$ с $f_2(t)$ равна свертке $f_2(t)$ с $f_1(t)$, т. е. $f_1(t) * f_2(t)=f_2(t) * f_1(t)$. Поэтому можно оставить $f_2(\tau)$ неизменной и графически определять свертку с помощью зеркального отображения $f_1(\tau)$.

СВЕРТКА ФУНКЦИИ С ЕДИНИЧНЫМ ИМПУЛЬСОМ

Свертка функции $f(t)$ с единичным импульсом дает саму функцию $f(t)$, что легко доказать на основании фильтрующего свойства (1.89):

$$f(t) * \delta(t)=\int_{-\infty}^{\infty} f(t)\,\delta(t-\tau)\,d\tau=f(t) \tag{1.130}$$

Это следует также из теоремы о свертке во времени и выражений $f(t)\longleftrightarrow F(\omega)$ и $\delta(t)\longleftrightarrow 1$. Поэтому $f(t) * \delta(t)\longleftrightarrow F(\omega)$. Следовательно, $f(t) * \delta(t)=f(t)$.

Этот результат очевиден графически. Так как единичный импульс сконцентрирован в одной точке и имеет единичную площадь, то интеграл свертки дает саму функцию $f(t)$. На основании ф-лы (1.130) можно записать:

$$f(t) * \delta(t-T) = f(t-T), \tag{1.131а}$$

$$f(t-t_1) * \delta(t-t_2) = f(t-t_1-t_2), \tag{1.131б}$$

$$\delta(t-t_1) * \delta(t-t_2) = \delta(t-t_1-t_2). \tag{1.131в}$$

Пример 1.12

Найдем графически свертку функции $f_1(t)$, изображенной на рис. 1.32*а* с парой единичных импульсов, каждый интенсивностью $к$ (рис. 1.32*б*). Согласно описанной выше процедуре графической свертки повернем $f_2(\tau)$ относительно оси координат, чтобы получить $f_2(-\tau)$. Так как $f_2(\tau)$ — четная функция τ, $f_2(-\tau)=f_2(\tau)$.

Из свойства единичного импульса воспроизводить в результате свертки саму функцию [ф-ла (1.130)] вытекает, что каждый импульс дает свертку в виде треугольного импульса высотою $Ак$ в начале координат ($t=0$). Следовательно, общая высота треугольного импульса в начале координат равна $2Ак$. Если функция $f_2(t-\tau)$ смещается в положительном направлении, импульс, первоначально расположенный в точке $-T$, перекрывается с треугольным импульсом вблизи $\tau=T$ и в результате воспроизводится треугольный импульс высотой $Ак$ при $t=2T$. Подобным же образом импульс, расположенный в точке T, воспроизводит треугольный импульс высотой $Ак$ при $t=-2T$. Окончательный результат свертки приведен на рис. 1.32*г*.

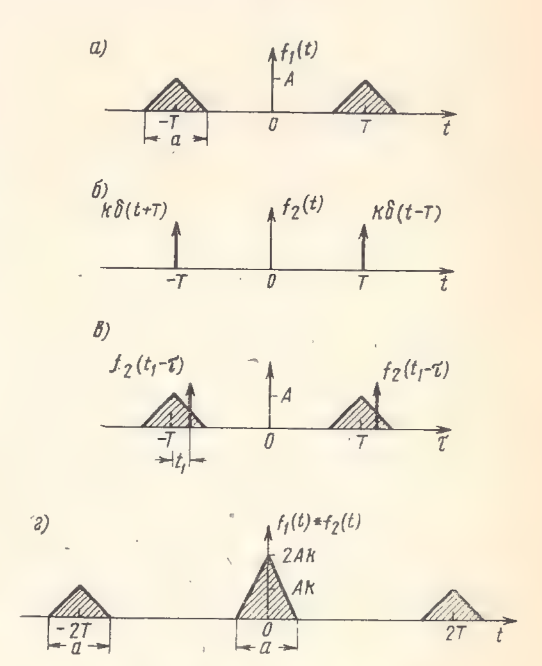


Рис. 1.32. Свертка функции с единичным импульсом:

а) функция $f_1(t)$; *б)* единичные импульсы [функция $f_2(t)$]; *в)* функции $f_1(t)$ и $f_2(t)$ при произвольном временно́м сдвиге t_1; *г)* график свертки функций $f_1(t)$ и $f_2(t)$

1.15. Дискретизация сигналов

ТЕОРЕМА ОТСЧЕТОВ ВО ВРЕМЕННОЙ ОБЛАСТИ

Теорема отсчетов занимает фундаментальное положение в теории связи. Формулируется она следующим образом.

Сигнал с ограниченным спектром, не имеющий спектральных составляющих с частотами выше f_m, *гц,* однозначно определяется

своими мгновенными значениями (отсчетами), разделенными одинаковыми интервалами времени, равными или меньше $1/(2f_m)$, *сек*.

Эта теорема известна как теорема *равномерных отсчетов*[1]). Она устанавливает, что если преобразование Фурье $f(t)$ равно нулю на частотах, выше некоторой частоты $\omega_m=2\pi f_m$, то в отсчетах, следующих равномерно через $1/(2f_m)$, содержится полная информация об $f(t)$ (рис. 1.33). Отсчеты функции осуществляются каждые T *сек* $[T<1/(2f_m)]$ или с частотой, большей или равной $2f_m$ отсчетов в секунду. Последовательные отсчеты обозначаются f_0, f_1, f_2 ... и т. д. Из рассмотренной теоремы следует, что эти отсчеты содержат информацию о сигнале $f(t)$ в любой момент времени. Однако частота отсчетов должна быть, по крайней мере, в два раза больше высшей частоты f_m, имеющейся в спектре сигнала $f(t)$. Иными словами, на каждый период наивысшей частотной составляющей сигнала должно приходиться, по крайней мере, по два отсчета.

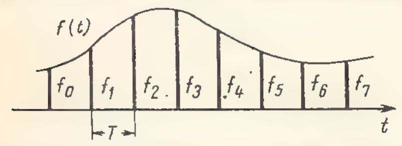


Рис. 1.33. Представление функции с помощью равномерных отсчетов

Теорему отсчетов можно легко доказать с помощью теоремы свертки в частотной области. Рассмотрим сигнал $f(t)$ с ограниченным спектром, не имеющий спектральных составляющих выше f_m, *гц*. Это означает, что $F(\omega)$ [преобразование Фурье сигнала $f(t)$] равно нулю при $|\omega|>\omega_m (\omega_m=2\pi f_m)$. Пусть функция $f(t)$ умножается на периодическую функцию $\delta_T(t)$ (рис. 1.34). Произведение указанных функций образует последовательность δ-импульсов, следующих через одинаковые промежутки времени T и имеющих интенсивности (т. е. площади), равные значениям $f(t)$ в соответствующие моменты времени. Обозначим эту импульсную функцию (выборку) (рис. 1.34*д*) через

$$f_s(t)=f(t)\,\delta_T(t). \tag{1.132}$$

Ранее было показано [ф-ла (1.106)], что преобразование Фурье периодической последовательности единичных импульсов есть также равномерная последовательность единичных импульсов $\omega_0\delta_{\omega_0}(\omega)$ (рис. 1.34*г*). Импульсы разделены равными частотными интервалами $\omega_0=2\pi/T$,

$$\delta_T(t) \longleftrightarrow \omega_0\delta_{\omega_0}(\omega).$$

[1]) Эта теорема является частным случаем общей теоремы отсчетов, которая формулируется следующим образом. Если сигнал имеет ограниченный спектр и если временно́й интервал делится на равные части, образующие подынтервалы, равные $T\leqslant 1/(2f_m)$, и если в каждом подынтервале в любой момент времени происходит отсчет сигнала, то мгновенные значения сигнала и моменты отсчетов в каждом подынтервале содержат всю информацию о первоначальном сигнале. См., например, H. S. Black, Modulation theory, Van Nostrand, New York, 1953, p. 41.

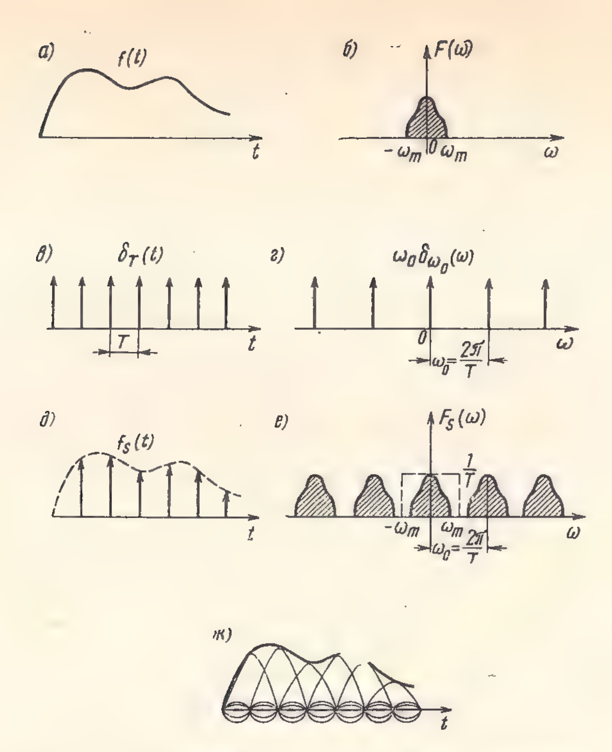


Рис. 1.34. Пояснение теоремы отсчетов с помощью свертки в частотной области:
а) непрерывная функция; *б)* ограниченный спектр функции $f(t)$; *в)* периодическая последовательность единичных импульсов; *г)* преобразование Фурье единичных импульсов; *д)* представление функции $f(t)$ последовательностью единичных импульсов; *е)* свертка спектров последовательности единичных импульсов и функции $f(t)$; *ж)* восстановление функции $f(t)$ с помощью фильтра нижних частот

Согласно теореме о свертке в частотной области преобразование Фурье произведения $f(t)\delta_T(t)$ находится как свертка $F(\omega)$ с $\omega_0\delta_{\omega_0}(\omega)$:

$$f_s(t) \longleftrightarrow \frac{1}{2\pi}\left[F(\omega) * \omega_0\delta_{\omega_0}(\omega)\right].$$

Подставив в это выражение $\omega_0 = 2\pi/T$, получим

$$f_s(t) \longleftrightarrow \frac{1}{T}\left[F(\omega) * \delta_{\omega_0}(\omega)\right]. \qquad (1.133)$$

Из ф-лы (1.133) следует, что спектр выборки $f_s(t)$ определяется сверткой $F(\omega)$ с последовательностью единичных импульсов.

Функции $F(\omega)$ и $\delta_{\omega_0}(\omega)$ (показанные на рис. 1.34*б* и *г* соответственно) можно свернуть графически методом, описанным в § 1.14.

Для этого функцию $\delta_{\omega_0}(\omega)$ нужно перевернуть относительно вертикальной оси $\omega=0$. Поскольку $\delta_{\omega_0}(\omega)$ — четная функция частоты ω, перевернутая функция совпадает с первоначальной функцией $\delta_{\omega_0}(\omega)$. Чтобы выполнить операцию свертки, всю последовательность импульсов $\delta_{\omega_0}(\omega)$ нужно сдвинуть в положительном направлении оси ω. При этом каждый импульс повторяет $F(\omega)$. Так как импульсы следуют через $\omega_0=2\pi/T$, операция свертки дает спектр $F(\omega)$, повторяющийся через каждые ω_0 *рад/сек*, как показано на рис. 1.34*е*. Итак, преобразование Фурье выборки $f_S(t)$ есть периодически повторяющийся через каждые ω_0 *рад/сек* спектр $F(\omega)$. Обозначим эту функцию через $F_s(\omega)$. Заметим, что $F(\omega)$ повторяется периодически без взаимного перекрытия при условии $\omega_0>2\omega_m$ или $2\pi/T>2(2\pi f_m)$. Это означает, что

$$T < 1/(2f_m). \tag{1.134}$$

Итак, если отсчеты $f(t)$ берутся через регулярные интервалы времени, не превышающие $1/(2f_m)$ *сек*, спектр $F_s(\omega)$ периодически повторяет $F(\omega)$ и поэтому содержит всю информацию о сигнале $f(t)$. Можно легко выделить $F(\omega)$ из $F_s(\omega)$, пропустив сигнал $f_S(t)$ через фильтр нижних частот, который подавляет все составляющие с частотами выше f_m. Характеристика идеального фильтра показана пунктирной линией на рис. 1.34*е*.

Заметим, что если интервал T превышает $1/(2f_m)$, то свертка $F(\omega)$ с $\delta_{\omega_0}(\omega)$ дает по-прежнему периодически повторяющуюся функцию $F(\omega)$. Но теперь соседние спектры перекрываются, и $F(\omega)$ нельзя без искажений выделить из $F_s(\omega)$. Максимальный интервал выборки $T=1/(2f_m)$ называется интервалом Найквиста[1]).

Показанную на рис. 1.34 свертку $F(\omega) * \delta_{\omega_0}(\omega)$ можно выполнить аналитически. Так как $\delta_{\omega_0}(\omega)=\sum_{n=-\infty}^{\infty}\delta(\omega-n\omega_0)$, то из ф-лы (1.133) следует, что

$$F_s(\omega)=\frac{1}{T}[F(\omega) * \delta_{\omega_0}(\omega)]=\frac{1}{T}\left[F(\omega) * \sum_{n=-\infty}^{\infty}\delta(\omega-\omega_0 n)=\right.$$

$$=\frac{1}{T}\sum_{n=-\infty}^{\infty}F(\omega) * \delta(\omega-n\omega_0).$$

Воспользовавшись соотношением (1.131а), получим

$$F_s(\omega)=\frac{1}{T}\sum_{n=-\infty}^{\infty}F(\omega-n\omega_0). \tag{1.135}$$

---

[1]) В отечественной литературе интервал $T=1/(2f_m)$ называется интервалом Котельникова. — *Прим. ред.*

Правая часть равенства (1.135) представляет собой периодически повторяющийся через ω_0 *рад/сек* спектр $F(\omega)$. Это полностью совпадает с результатом, полученным графической сверткой.

ВОССТАНОВЛЕНИЕ СИГНАЛА $f(t)$ ПО ЕГО ОТСЧЕТАМ

Как было выяснено ранее, исходную функцию можно восстановить, пропустив последовательность отсчетов через фильтр нижних частот с частотой среза ω_m. В частотной области эта операция очевидна. Попробуем теперь восстановить функцию $f(t)$ по ее отсчетам во временно́й области.

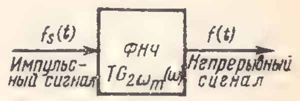


Рис. 1.35. Схема восстановления непрерывной функции по ее отсчетам

Рассмотрим сигнал $f(t)$, отсчеты которого следуют с минимально допустимой частотой ($2f_m$ отсчетов в секунду). В этом случае $T=1/(2f_m)$, $\omega_0=2\pi/T=4\pi f_m=2\omega_m$. Поэтому ф-ла (1.135) принимает вид

$$F_s(\omega)=\frac{1}{T}\sum_{n=-\infty}^{\infty}F(\omega-2n\,\omega_m). \tag{1.136}$$

Очевидно, операция фильтрации эквивалентна умножению $F_s(\omega)$ на стробирующую функцию $G_{2\omega_m}(\omega)$. Следовательно, из ф-лы (1.136) вытекает, что

$$F_s(\omega)\,G_{2\omega_m}(\omega)=\frac{1}{T}F(\omega),$$

$$F(\omega)=TF_s(\omega)\,G_{2\omega_m}(\omega). \tag{1.137}$$

Операция восстановления сигнала $f(t)$ фильтром нижних частот поясняется рис. 1.35. Передаточная функция фильтра

$$H(\omega)=TG_{2\omega_m}(\omega)=\frac{1}{2f_m}G_{2\omega_m}(\omega).$$

Применив к выражению (1.137) теорему о временно́й свертке, получим

$$f(t)=Tf_s(t)*\frac{\omega_m}{\pi}Sa(\omega_m t)=f_s(t)*Sa(\omega_m t). \tag{1.138}$$

Импульсная функция $f_s(t)=\sum_n f_n\delta(t-nT)$, где f_n представляет собой n-й отсчет $f(t)$. Следовательно,

$$f(t)=\sum_n f_n\,\delta(t-nT)*Sa(\omega_m t)=\sum_n f_n Sa[\omega_m(t-nT)]= \tag{1.139}$$

$$=\sum_n f_n Sa(\omega_m t-n\pi). \tag{1.140}$$

Формула (1.140) показывает, что для восстановления сигнала $f(t)$ во временно́й области каждый отсчет умножается на функцию отсчетов и все полученные колебания суммируются. Этот процесс иллюстрируется рис. 1.34*ж*.

Большинство практических сигналов хорошо аппроксимируется сигналами с ограниченным спектром. Следует подчеркнуть, что, строго говоря, сигналов с ограниченным спектром не существует. Можно показать, что, если сигнал существует на конечном интервале времени, он содержит составляющие всех частот[1]). Однако для всех практических сигналов спектральная плотность на высоких частотах ничтожно мала. Бо́льшая часть энергии переносится составляющими в определенном частотном интервале, и поэтому практически сигнал можно рассматривать как имеющий ограниченный спектр. Ошибка, возникающая при отбрасывании высших частотных составляющих, пренебрежимо мала.

Теорема отсчетов является весьма важной, так как она позволяет заменить непрерывный сигнал с ограниченным спектром дискретной последовательностью его отсчетов без какой-либо потери информации. Таким образом, информационное содержание непрерывного сигнала с ограниченным спектром эквивалентно дискретным частям информации. Поскольку теорема отсчетов определяет наименьшее число отсчетов, необходимое для воспроизведения непрерывного сигнала, проблема передачи такого сигнала сводится к передаче конечного числа импульсов[2]). Такая дискретная информация может быть передана группой импульсов, амплитуды которых пропорциональны значениям отсчетов (амплитудно-импульсная модуляция). Другими видами модуляции являются фазово-импульсная модуляция, при которой изменяется положение импульсов во времени, широтно-импульсная модуляция (изменяется длительность импульсов пропорционально значениям отсчетов выборок), а также кодово-импульсная модуляция (отсчеты представляются кодом, образованным из групп импульсов[3]).

ТЕОРЕМА ОТСЧЕТОВ В ЧАСТОТНОЙ ОБЛАСТИ

Теорема отсчетов во временно́й области имеет эквивалент в частотной области, который устанавливает, что ограниченный во времени сигнал, равный нулю при $|t|>T$, полностью определяется

---

[1]) Это следует из критерия Пэйли-Винера (§ 2.5). Если функция $F(\omega)$ ограничена по частоте [$F(\omega)=0$ при $(\omega)>\omega_m$], то она удовлетворяет условию Пэйли-Винера и, следовательно, ее обратное преобразование $f(t)$ существует на всей отрицательной полуоси времени. Поэтому ограниченный по спектру сигнал существует на всем бесконечном интервале времени, и наоборот, ограниченный во времени сигнал не может быть ограниченным по спектру.

[2]) Импульсный сигнал дискретен во времени; амплитуды импульсов могут принимать любые непрерывные значения в пределах от **максимального до минимального** значений $f(t)$. — *Прим. ред.*

[3]) Кодово-импульсная модуляция требует, помимо временно́й, еще и дискретизацию по значениям отсчетов — квантование. — *Прим. ред.*

отсчетами своего частотного спектра, взятыми через одинаковые интервалы, равные или меньшие $1/(2T)$ *гц* (или π/T *рад/сек*):

$$F(\omega)=\sum_{n=-\infty}^{\infty} F\left(\frac{n\pi}{T}\right) Sa(\omega T-n\pi). \qquad (1.141)$$

Доказательство этой теоремы подобно приведенному выше для временно́й области, в котором $F(\omega)$ и $f(t)$ поменялись ролями.

ЗАДАЧИ.

1.1. Покажите, что прямоугольная функция, изображенная на рис. 1.3, на интервале $(0,2\pi)$ ортогональна сигналам $\cos t$, $\cos 2t$, ..., $\cos nt$ для всех целых n, т. е. что эта функция не имеет в своем составе функций типа $\cos nt$ (n—целое число).

1.2. Покажите, что если два сигнала $f_1(t)$ и $f_2(t)$ ортогональны на интервале (t_1, t_2), то энергия сигнала $[f_1(t)+f_2(t)]$ равна сумме энергий $f_1(t)$ и $f_2(t)$. Энергия сигнала $f(t)$ на интервале (t_1, t_2) определяется как $\int_{t_1}^{t_2} f^2(t)dt$.

Распространите этот результат на n взаимоортогональных сигналов.

1.3. Прямоугольная функция $f(t)$ на рис. 1.3 аппроксимирована сигналом $\frac{4}{\pi}\sin t$. Покажите, что функция ошибки $f_e=f(t)-\frac{4}{\pi}\sin t$ ортогональна функции $\sin t$ на интервале $(0,2\pi)$. Дайте этому физическое объяснение. Покажите теперь, что энергия сигнала $f(t)$ равна сумме энергий $f_e(t)$ и $\frac{4}{\pi}\sin t$.

1.4. Аппроксимируйте двумя первыми ненулевыми членами полинома Лежандра прямоугольную функцию, изображенную на рис. 1.3. Найдите среднюю квадратическую ошибку аппроксимации, если последняя выполнена: а) первым членом полинома и б) первым и вторым членами полинома. Сравните эту аппроксимацию с аппроксимацией синусоидальными функциями [ф-ла (1.38)].

1.5. Докажите, что если $f_1(t)$ и $f_2(t)$ являются комплексными функциями действительной переменной t, то составляющая $f_2(t)$, содержащаяся в $f_1(t)$ на интервале (t_1, t_2), определяется выражением

$$C_{12}=\frac{\int_{t_1}^{t_2} f_1(t)\, f_2^*(t)\, dt}{\int_{t_1}^{t_2} f_2(t)\, f_2^*(t)\, dt}.$$

Составляющая определяется в обычном смысле критерия минимальной величины средней квадратической ошибки. Покажите теперь, что сигналы $[f_1(t)-C_{12}f_2(t)]$ и $f_2(t)$ являются взаимно ортогональными.

1.6. Найдите составляющую $\sin\omega_2 t$, содержащуюся в другой функции $\sin\omega_1 t$ на интервале $(-T, T)$ для всех действительных значений ω_1 и ω_2 $(\omega_1\neq\omega_2)$. Как изменяется величина этой составляющей при изменении периода T? Докажите, что при $T\to\infty$ составляющая $\sin\omega_2 t$ в функции $\sin\omega_1 t$ стремится к нулю. Распространите этот результат на любую пару функций: $\sin\omega_1 t$, $\sin\omega_2 t$, $\cos\omega_1 t$, $\cos\omega_2 t$.

1.7. Две периодические функции $f_1(t)$ и $f_2(t)$ с периодами T и $\sqrt{2}\,T$ соответственно имеют нулевые постоянные составляющие. Покажите, что на интервале $(-\infty<t<\infty)$ составляющая функции $f_2(t)$ не содержится в $f_1(t)$. Докажите, что это справедливо для любых двух периодических функций, если отношение их периодов равно иррациональному числу и при условии, что либо $f_1(t)$, либо $f_2(t)$, или обе функции имеют нулевое среднее (нулевую постоянную составляющую).

1.8. Определите, являются ли приведенные ниже функции периодическими или непериодическими. Для периодической функции найдите ее период.

а) $a \sin t + b \sin 2t$, д) $a \cos t + b \sin \sqrt{2}\, t$,

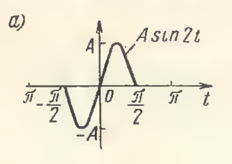
б) $a \sin 5t + b \cos 8t$, е) $a \sin (3t/2) + b \cos (16t/15) + c \sin (t/29)$,

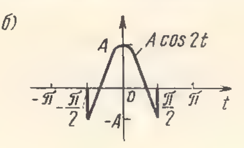
в) $a \sin 2t + b \cos \pi t$, ж) $(a \sin t)^3$,

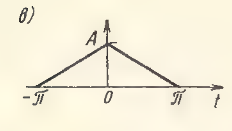
г) $a \cos 2t + b \sin 7t + c \sin 13t$, з) $(a \sin 2t + b \sin 5t)^2$.

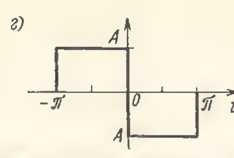
1.9. Представьте функции e^t и t^2 на интервале $(0<t<1)$ тригонометрическим и экспоненциальным рядами.

1.10. Каждую из четырех функций, изображенных на рисунке, необходимо представить тригонометрическим рядом Фурье на интервале $(-\pi, \pi)$. Аппроксимируйте затем эти функции конечным числом членов ряда, выбрав его так, чтобы ошибка аппроксимации не превышала 1% от полной энергии $f(t)$.









1.11. Если периодический сигнал удовлетворяет определенным условиям симметрии, то вычисление коэффициентов ряда Фурье несколько упрощается. Покажите справедливость следующих утверждений:

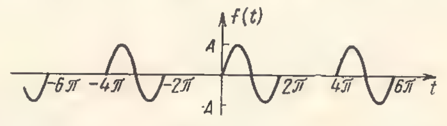
а) если $f(t)=f(-t)$ (четная симметрия), все синусоидальные члены тригонометрического ряда Фурье отсутствуют;

б) если $f(t)=-f(-t)$ (нечетная симметрия), все косинусоидальные члены тригонометрического ряда Фурье отсутствуют;

в) если $f(t)=-f(t+T/2)$ (симметрия поворота), все четные гармоники отсутствуют.

Покажите, что в каждом случае коэффициенты Фурье можно вычислить интегрированием периодического сигнала только в течение полупериода.

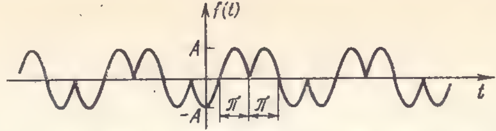
1.12. Периодическая функция образована в результате удаления ряда периодов синусоидального колебания (см. рисунок). Напишите тригонометрический или экспоненциальный ряд Фурье путем непосредственного вычисления его коэффициентов.



Если $f(t)$ сдвинуть на π *сек* влево, новая функция $f(t+\pi)$ будет нечетной функцией времени, преобразование Фурье которой содержит только синусоидальные члены. Напишите ряд Фурье для функции $f(t+\pi)$ и из него получите ряд Фурье для $f(t)$. Повторите указанное, сдвинув $f(t)$ вправо на π *сек*.

1.13. Периодическая функция $f(t)$ образована из синусоиды в результате поворота относительно горизонтальной оси некоторых периодов, как это показано на рисунке. Непосредственным вычислением определите коэффициенты ряда Фурье и запишите ряд.

Если $f(t)$ сдвинуть вправо на $\pi/2$ *сек*, новая функция будет четной функцией времени, преобразование Фурье которой содержит только косинусоидальные члены. Определите ряд Фурье для функции $f(t-\pi/2)$ и из него получите ряд Фурье для $f(t)$.

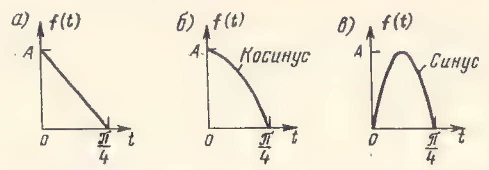


Функция $f(t+\pi/2)$ является нечетной функцией времени. Напишите для нее ряд Фурье и из него получите ряд Фурье для $f(t)$.

Данную функцию можно также выразить через колебание, рассмотренное в задаче 1.12. Используя результаты последней, напишите ряд Фурье для $f(t)$.

1.14. Каждую из функций $f(t)$, показанных на рисунке, разложите в тригонометрический ряд Фурье, используя следующие составляющие:

а) синусоидальные и косинусоидальные с частотами $\omega=4$, 8, 12, 16 и т. д., а также постоянную составляющую;



б) только синусоидальные с частотами $\omega=2$, 6, 10, 14 и т. д.;

в) постоянную и косинусоидальные составляющие с частотами $\omega=8/3$, 16/3, 8, 32/3 ... и т. д.;

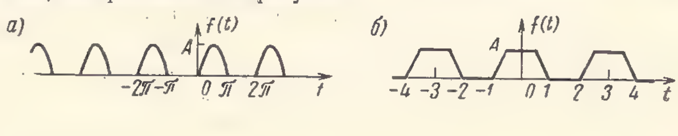
г) постоянную и косинусоидальные составляющие с частотами $\omega=2$, 4, 6, 8 и т. д.;

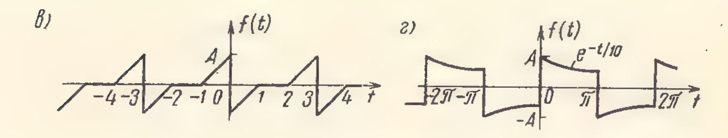
д) постоянную составляющую, а также синусоидальные и косинусоидальные с частотами $\omega=1$, 2, 3, 4 и т. д.;

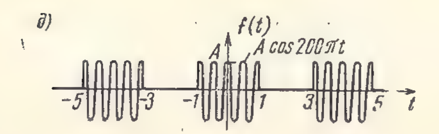
е) только синусоидальные с частотами $\omega=1$, 3, 5, 7 и т. д.

Если функция $f(t)$ аппроксимируется конечным числом членов вышеуказанных рядов, то какой ряд Вы выберете? Приведите качественное и аналитическое обоснование.

1.15. Найдите ряд Фурье и изобразите частотный спектр для периодических функций, изображенных на рисунке.







1.16. Покажите, что произвольную функцию $f(t)$ можно представить суммой четной $f_e(t)$ и нечетной $f_o(t)$ функций: $f(t)=f_e(t)+f_o(t)$. Найдите четные и нечетные составляющие функций: $u_1(t)$, $e^{-at}u_1(t)$, e^{it}.

Указание. $f_e(t)=\frac{1}{2}[f(t)+f(-t)]$, $f_o(t)=\frac{1}{2}[f(t)-f(-t)]$.

1.17. Покажите, что для четной периодической функции коэффициенты ряда Фурье — действительные, а для нечетной — мнимые.

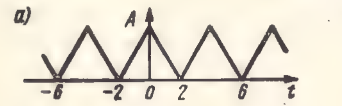
1.18. Ряд Фурье непрерывной периодической функции $f(t)$ определяется выражением (1.55). Покажите, что функция df/dt также является периодической с тем же периодом и может быть представлена рядом

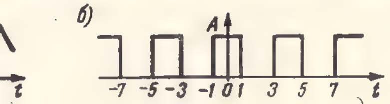
$$\frac{df}{dt}=\sum_{n=-\infty}^{\infty}(\mathrm{i}\,n\,\omega_0\,F_n)\,\mathrm{e}^{\mathrm{i}\,n\,\omega_0\,t}.$$

Покажите, далее, что для функции $f(t)$, не имеющей постоянной составляющей (т. е. $F_0=0$), интеграл ее также является периодической функцией и может быть представлен рядом

$$\int f(t)\,dt=\sum_{n=-\infty}^{\infty}\frac{F_n}{\mathrm{i}\,n\,\omega_0}\,\mathrm{e}^{\mathrm{i}\,n\,\omega_0\,t}.$$

1.19. Представьте ряд Фурье для периодических функций, показанных на





рисунке. Как изменяются коэффициенты F_n с изменением n? Как можно физически объяснить полученные результаты, используя задачу 1.18?

1.20. Периодическая функция $f(t)$ определяется только первыми n гармониками; все остальные равны нулю. Такой сигнал называют сигналом с ограниченным спектром. Покажите, что периодический сигнал с ограниченным спектром однозначно определяется своими значениями в любых $(2n+1)$ точках за период.

1.21. Покажите, что преобразование Фурье сигнала $f(t)$ можно выразить в виде

$$\mathcal{F}\,[f\,(t)]=F\,(\omega)=\int_{-\infty}^{\infty}f\,(t)\cos\omega\,t\,dt-\mathrm{i}\int_{-\infty}^{\infty}f\,(t)\sin\omega\,t\,dt.$$

Покажите также, что если $f(t)$ является четной функцией t, то $F(\omega)=2\int_{0}^{\infty}f(t)\cos\omega t\,dt$, а если нечетной, то $F(\omega)=-2\mathrm{i}\int_{0}^{\infty}f(t)\sin\omega t\,dt$.

Докажите таким образом, что между функциями $f(t)$ и $F(\omega)$ существует следующая взаимосвязь:

$f\,(t)$	$F\,(\omega)$
Действительная и нечетная	Мнимая и нечетная
Действительная и четная	Действительная и четная
Мнимая и четная	Мнимая и четная
Комплексная и четная	Комплексная и четная
Комплексная и нечетная	Комплексная и нечетная

1.22. Функцию $f(t)$ можно выразить суммой четной и нечетной функций (см. задачу 1.16): $f(t)=f_e(t)+f_o(t)$, где $f_e(t)$ — четная, а $f_o(t)$ — нечетная функции $f(t)$.

Покажите, что если $F(\omega)$ является преобразованием Фурье действительного сигнала $f(t)$, то $\mathrm{Re}[F(\omega)]$ будет преобразованием Фурье $f_e(t)$, а $\mathrm{Im}[F(\omega)]$ — преобразованием Фурье $f_o(t)$.

Покажите, что если $f(t)$ — комплексная функция $f(t)=f_r(t)+\mathrm{i}f_i(t)$ и $F(\omega)$ — ее преобразование Фурье, то

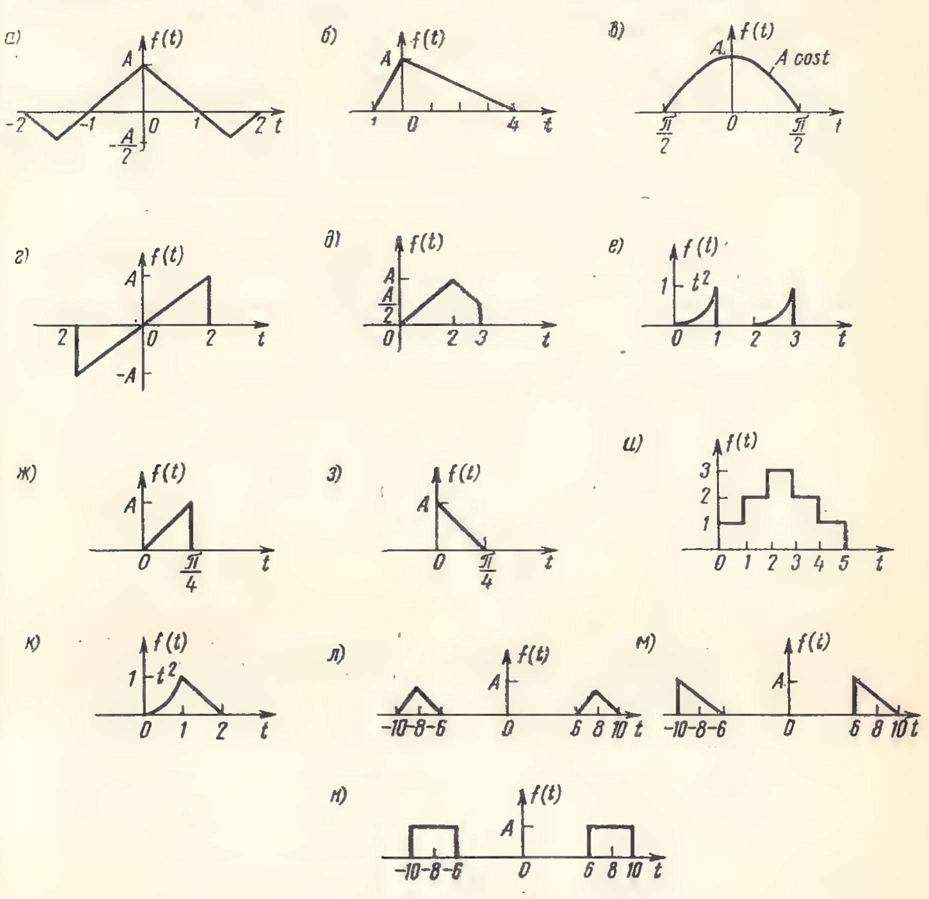
$$\mathscr{F}[f_r(t)]=\frac{1}{2}[F(\omega)+F^*(-\omega)],$$

$$\mathscr{F}[f_i(t)]=\frac{1}{2}[F(\omega)-F^*(-\omega)].$$

Указание. $f^*(t)=f_r(t)-\mathrm{i}f_i(t)$, $F[f^*(t)]=F^*(-\omega)$.

1.23. Найдите преобразование Фурье функций $f(t)$, изображенных на рисунке.

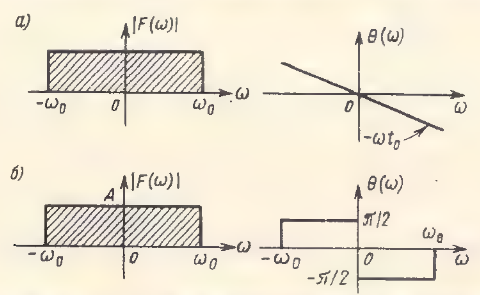
1.24. Выведите соотношения (1.100а) и (1.100б), используя теорему о модуляции и ф-лу (1.96).



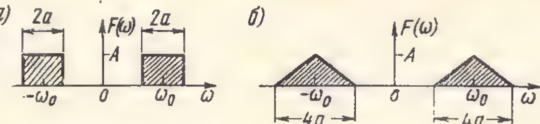
1.25. Вычислите на основании фильтрующего свойства единичного импульса следующие интегралы:

$$\int_{-\infty}^{\infty} \delta(t-2)\sin t\,dt, \qquad \int_{-\infty}^{\infty} \delta(t+3)\,e^{-t}\,dt, \qquad \int_{-\infty}^{\infty} \delta(1-t)(t^3+4)\,dt.$$

1.26. Определите функции $f(t)$, преобразования Фурье которых показаны на рисунке.

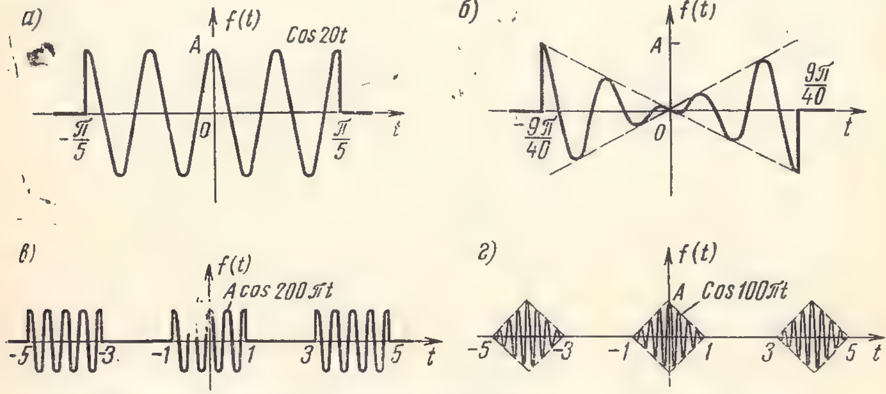


1.27. Найдите функцию $f(t)$, преобразование Фурье которой показано на рисунке.



Указание. Используйте теорему о модуляции.

1.28. Найдите преобразования Фурье функций, показанных на рисунке, используя теорему о модуляции. Изобразите для каждого случая частотный спектр.



Примечание: на рисунке (б) показана функция

$$f(t) = \frac{-40A}{9\pi}\, t\cos 20t \quad \text{при} \quad t < |9\pi/40|$$

1.29. Определите преобразования Фурье следующих функций, если $f(t) \longleftrightarrow F(\omega)$: а) $t f(2t)$, б) $(t-2)f(t)$, в) $(t-2)f(2-t)$, г) $t\dfrac{df}{dt}$, д) $f(1-t)$, е) $(1-t)f(1-t)$.

1.30. Найдите преобразования Фурье функций $f(t)$, показанных в задаче 1.23, используя теорему о дифференцировании в частотной области, теорему о временно́м сдвиге и табл. 1.1.

1.31. Найдите преобразования Фурье функций $f(t)$ (рис. к задаче 1.23 л, м, н), используя ф-лу (1.118).

1.32. Найдите преобразование Фурье функции, показанной на рисунке:

а) непосредственным интегрированием,

б) используя только свойство интегрирования во временно́й области и табл. 1.1,

в) используя только свойства дифференцирования во временно́й области, теорему о временном сдвиге и табл. 1.1,

г) используя только теорему о временно́м сдвиге и табл. 1.1.

A
-τ
τ
t
-A

1.33. Момент n-го порядка функции $f(t)$ определяется формулой $m_n = \int_{-\infty}^{\infty} t^n f(t)dt$. Используя теорему о дифференцировании в частотной области, покажите, что

$$m_n = (\mathrm{i})^n \frac{d^n F(0)}{d\,\omega^n}.$$

На основании этого результата покажите, что разложение $F(\omega)$ в ряд Тейлора имеет вид

$$F(\omega) = m_0 - \mathrm{i}\, m_1\, \omega - m_2\, \omega^2/2! + \mathrm{i}\, m_3\, \omega^3/3! + m_4\, \omega^4/4! + \ldots =$$

$$= \sum_{n=0}^{\infty} (-\mathrm{i})^n\, m_n \frac{\omega^n}{n!}.$$

Определите различные моменты стробирующей функции и на основании последней формулы найдите ее преобразование Фурье.

1.34. Покажите, что если $f(t) \longleftrightarrow F(\omega)$, то

$$|F(\omega)| < \int_{-\infty}^{\infty} |f(t)|\, dt; \quad |F(\omega)| < \frac{1}{|\omega|} \int_{-\infty}^{\infty} \left|\frac{df}{dt}\right| dt;$$

$$|F(\omega)| < \frac{1}{\omega^2} \int_{-\infty}^{\infty} \left|\frac{d^2 f}{dt^2}\right| dt.$$

Эти неравенства определяют верхние границы $|F(\omega)|$.

1.35. Вычислите следующие интегралы свертки:

а) $u_1(t) * u_1(t)$, б) $u_1(t) * \mathrm{e}^{-t}\, u_1(t)$, в) $\mathrm{e}^{-t}\, u_1(t) * \mathrm{e}^{-2t}\, u_1(t)$,

г) $u_1(t) * t\, u_1(t)$, д) $\mathrm{e}^{-t}\, u_1(t) * t\, u_1(t)$, е) $\mathrm{e}^{-2t}\, u_1(t) * \mathrm{e}^{-t}$.

Проверьте полученные результаты от (б) до (е), используя преобразование Фурье.

1.36. Покажите, что если $f(t)$ является непрерывным сигналом, ограниченным по спектру частотой ω_m *рад сек*, то

$$\frac{\kappa}{\pi}\,[f(t) * Sa(\kappa t)] = f(t) \quad \text{для} \quad \kappa \geqslant \omega_m.$$

Покажите, далее, что

$$\frac{\omega_n}{\pi}\,[Sa\,(\omega_m\,t) * Sa\,(\omega_n\,t)] = Sa\,(\omega_m\,t) \quad \text{для} \quad \omega_n \geqslant \omega_m.$$

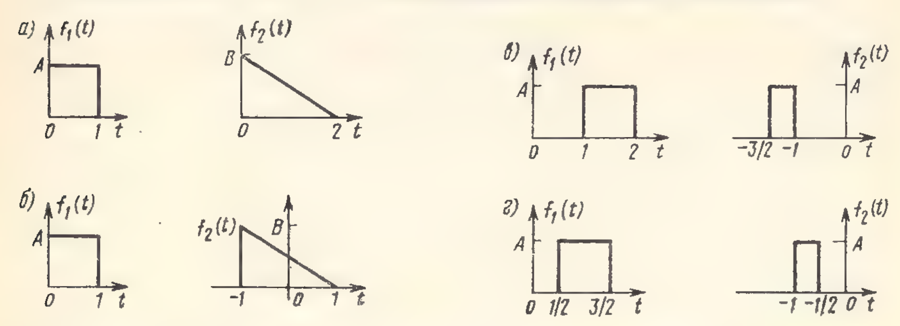
1.37. Определите обратное преобразование Фурье функции $Sa^2(Wt)$, используя теорему временно́й свертки. Вычислите графически интеграл свертки.

1.38. Определите минимальную частоту отсчетов и интервал Котельникова для следующих сигналов:

а) $Sa\,(100\,t)$, в) $Sa\,(100\,t) + Sa\,(50\,t)$,

б) $Sa^2\,(100\,t)$, г) $Sa\,(100\,t) + Sa^2\,(60\,t)$.

1.39. Вычислите $f_1 * f_2$ и $f_2 * f_1$ для функций, показанных на рисунке.

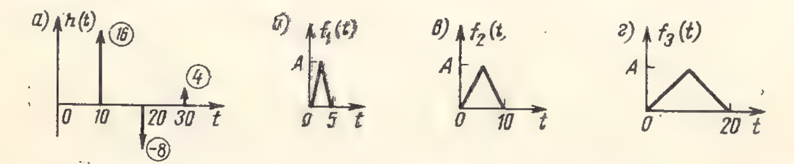


1.40. Покажите, что причинную (каузальную) функцию $f(t)$ (т. е. функцию, которая равна нулю для $t<0$) можно выразить непрерывной суммой единичных наклонных функций:

$$f\,(t) = \int\limits_0^t \frac{d^2 f}{d\,\tau^2}\,(t-\tau)\,d\,\tau.$$

Указание. Используйте теорему свертки.

1.41. Вычислите свертки $h(t) * f_1(t)$, $h(t) * f_2(t)$ и $h(t) * f_3(t)$ функций, показанных на рисунке.



1.42. Теорема об интегрировании во временно́й области (1.119б) справедлива в том случае, если $\int\limits_{-\infty}^{\infty} f(t)dt = 0$.

Покажите, что если это условие не удовлетворяется, то

$$\int\limits_{-\infty}^{t} f\,(\tau)\,d\,\tau \longleftrightarrow \pi\,F\,(\omega)\,\delta\,(\omega) + \frac{1}{i\,\omega}\,F\,(\omega).$$

Указание. Выразите $\int_{-\infty}^{t} f(\tau)d\tau$ как свертку функций $f(t)$ и $u_1(t)$.

1.43. Пусть сигнал $f(t)$, показанный на рис. 1.14, имеет преобразование Фурье $F(\omega)$. Покажите, что преобразование Фурье периодической функции $f_T(t)$, образованной повторением $f(t)$ через интервал T *сек* (рис. 1.15), имеет вид

$$f_T(t) \longleftrightarrow \frac{2\pi}{T} F(\omega)\, \delta_{\omega_0}(\omega) \quad \left(\omega_0 = \frac{2\pi}{T}\right).$$

Указание. Разложите $f_T(t)$ в ряд Фурье и заметьте, что коэффициент F_n при n-м члене этого ряда равен $(1/T)F(\omega)\,|_{\omega = n\omega_0}$; или выразите $f_T(t)$ в виде свертки $f(t)$ и $\delta_T(t)$.

Найдите преобразование Фурье стробирующей функции $G_\tau(t)$ для $\tau = 1/20$ *сек*. На основании результата, полученного ранее, изобразите преобразование Фурье периодической стробирующей функции с периодом 1/4 *сек*. Сравните результаты с полученными в примере 1.7 этой главы.

1.44. Пусть изображенная на рис. 1.14 функция $f(t)$ имеет преобразование Фурье $F(\omega)$. Образуем новую функцию, состоящую из n периодов по T *сек*. Полученная функция $f_n(t)$ существует на интервале $(-nT/2 < t < nT/2)$ и равна нулю вне этого интервала. Покажите, что

$$f_n(t) \longleftrightarrow F(\omega)\, \frac{\sin(n\omega T/2)}{\sin(\omega T/2)}.$$

Изобразите функцию $[\sin(n\omega T/2)]/[\sin(\omega T/2)]$ для $n = 15$. Покажите, что при $n \to \infty$ эта функция стремится к последовательности единичных импульсов, т. е.

$$\lim_{n\to\infty} \frac{\sin(n\omega T/2)}{\sin(\omega T/2)} = \omega_0\, \delta_{\omega_0}(\omega) \quad \left(\omega_0 = \frac{2\pi}{T}\right).$$

Указание. Используйте выражение

$$\sum_{k=-m}^{m} e^{ikx} = \frac{\sin(nx/2)}{\sin(x/2)} \quad (n = 2m + 1).$$

1.45. Функция $h(t) \longleftrightarrow H(\omega)$ задана так, что она равна нулю для всех $t < 0$ (такая функция называется причинной). Покажите, что если $R(\omega)$ и $X(\omega)$ — действительная и мнимая части $H(\omega)$ и $h(t)$ не содержит единичный импульс в начале координат, то

$$R(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{X(y)}{\omega - y}\, dy, \quad X(\omega) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{R(y)}{\omega - y}\, dy.$$

Эти уравнения называются парой преобразований Гильберта. Распространите полученный результат на случай, когда $h(t) = 0$ для всех $t > 0$.

Указание. Выразите $h(t)$ в виде четной и нечетной составляющих $h_e(t)$ и $h_o(t)$ (см. задачу 1.16). Из результатов задачи 1.22 следует, что $h_e(t) \longleftrightarrow R(\omega)$ и $h_o(t) \longleftrightarrow iX(\omega)$. Заметим также, что причинная функция имеет составляющие

$$h_e(t) = h_o(t) \operatorname{sign} t \quad \text{и} \quad h_o(t) = h_e(t) \operatorname{sign} t.$$

Теперь используйте ф-лу (1.95) и теорему о свертке.

1.46. Покажите, что $\dfrac{d^n\delta(t)}{dt^n} \longleftrightarrow (i\omega)^n$ и $t^n \longleftrightarrow 2\pi i^n \dfrac{d^n\delta(\omega)}{d\omega^n}$.

ГЛАВА

2

Передача сигналов и спектры плотности мощности

2.1. Передача сигналов через линейные системы

Характерной особенностью линейных систем является справедливость для них принципа суперпозиции. Это означает, что если $r_1(t)$ — реакция (отклик) системы на воздействие $f_1(t)$, а $r_2(t)$ — реакция на другое воздействие $f_2(t)$, то реакция системы на сложное воздействие $f_1(t)+f_2(t)$ равна $r_1(t)+r_2(t)$. В общем виде принцип суперпозиции формулируется следующим образом: реакция на воздействие $\alpha f_1(t)+\beta f_2(t)$ равна $\alpha r_1(t)+\beta r_2(t)$ для произвольных постоянных α и β.

Использование принципа суперпозиции облегчает определение реакции линейной системы на заданное воздействие. Последнее можно выразить суммой простых функций, для каждой из которых реакция системы легко находится. В гл. 1 показано, что произвольная функция $f(t)$ может быть выражена с помощью преобразования Фурье как непрерывная сумма экспоненциальных функций. Это представление можно применить для определения отклика системы методом интеграла Фурье (или интеграла Лапласа). В данном случае воспользуемся другим классом элементарных функций, а именно, единичными импульсами. Сначала выразим сигнал $f(t)$ как непрерывную сумму единичных импульсов, для чего воспользуемся ф-лой (1.130):

$$f(t)=f(t)*\delta(t)=\int_{-\infty}^{\infty} f(\tau)\,\delta(t-\tau)\,d\tau.$$

Правая часть этой формулы представляет собой непрерывную сумму (интеграл) единичных импульсов. Действительно, записав интеграл как предельную форму дискретной суммы, получим

$$f(t)=\lim_{\Delta\tau\to 0}\sum_{\tau=-\infty}^{\infty}[f(\tau)\,\Delta\tau]\,\delta(t-\tau). \qquad (2.1)$$

Здесь $f(t)$ выражается суммой единичных импульсов. Действующий в момент $t=\tau$ импульс имеет интенсивность (площадь) $f(\tau)\Delta\tau$.

Если $h(t)$ — реакция системы на единичный импульс $\delta(t)$, то реакция системы на $[f(\tau)\Delta\tau]\delta(t-\tau)$ равна[1] $f(t)\Delta\tau h(t-\tau)$, а полный отклик $r(t)$ на воздействие $f(t)$

$$r(t)=\lim_{\Delta\tau\to 0}\sum_{\tau=-\infty}^{\infty}[f(t)\Delta\tau]\,h(t-\tau)=\int_{-\infty}^{\infty}f(\tau)\,h(t-\tau)\,d\tau= \tag{2.2}$$

$$=f(t)*h(t). \tag{2.3}$$

Применив теорему временно́й свертки, найдем

$$R(\omega)=F(\omega)\,H(\omega), \tag{2.4}$$

где

$$r(t)\longleftrightarrow R(\omega),\quad f(t)\longleftrightarrow F(\omega)\text{ и } h(t)\longleftrightarrow H(\omega).$$

Функция $H(\omega)$ называется передаточной функцией (коэффициентом передачи) системы[2].

Если сигнал $f(t)$ начинается в момент $t=0$ и равен нулю при $t<0$, то нижний предел в ф-ле (2.2) можно заменить на нуль. Если, кроме того, $h(t)=0$ для $t<0$ (это условие справедливо для всех физических реализуемых систем), то $h(t-\tau)=0$ для $\tau>t$. Следовательно, верхний предел в интеграле (2.2) можно заменить на t:

$$r(t)=\int_0^t f(\tau)\,h(t-\tau)\,d\tau. \tag{2.5}$$

Заметим, что ф-ла (2.2) — общая, тогда как ф-лы (2.4) и (2.5) представляет собой частные случаи $f(t)=0$ и $h(t)=0$ при $t<0$.

2.2. Фильтрующие свойства линейных систем

Действующий на входе заданной системы сигнал $f(t)$ преобразуется ею в выходной сигнал (реакцию) $r(t)$ вполне определенным образом, следовательно, это преобразование можно рассматривать как характеристику системы. Если $F(\omega)$ — функция спектральной плотности входного сигнала, то функция спектральной плотности реакции равна $F(\omega)\,H(\omega)$, т. е. система изменяет спектральную плотность входного сигнала, действуя как фильтр, изменяющий частотные составляющие. Одни частотные составляющие увеличи-

[1]) Здесь подразумевается, что система линейна и инвариантна во времени. В этих системах параметры не изменяются во времени, поэтому, если $h(t)$ является реакцией системы на $\delta(t)$, то $h(t-\tau)$ будет реакцией на $\delta(t-\tau)$. Существует класс линейных систем (линейных систем с переменными параметрами), в которых один или более параметров могут изменяться во времени, и реакция на воздействие $\delta(t-\tau)$ может быть не равна $h(t-\tau)$.

[2]) Можно показать, что реакция системы на воздействие $e^{i\omega t}$ равна $H(\omega)e^{i\omega t}$.

ваются по амплитуде, другие уменьшаются, некоторые могут оставаться без изменений. Подобным же образом изменяются фазовые сдвиги каждой частотной составляющей сигнала при прохождении его через систему. Таким образом, передаточная функция $H(\omega)$ представляет собой реакцию системы на составляющие разных частот. Поэтому $H(\omega)$ действует на различные частотные составляющие как весовая функция и получающийся отклик имеет спектральную плотность $F(\omega)H(\omega)$ (рис. 2.1).

$f(t)$, $F(\omega)$ → $h(t)$, $H(\omega)$ → $r(t)$, $F(\omega)H(\omega)$

Рис. 2.1. К пояснению прохождения сигнала через линейную систему

Рассмотрим простейшую RC цепь, показанную на рис. 2.2*а*, на вход aa' которой поступает прямоугольный импульс, приведенный на рис. 2.2*в*. Откликом является напряжение $u_0(t)$, наблюдаемое на выходе цепи bb'. Здесь же показана функция спектральной

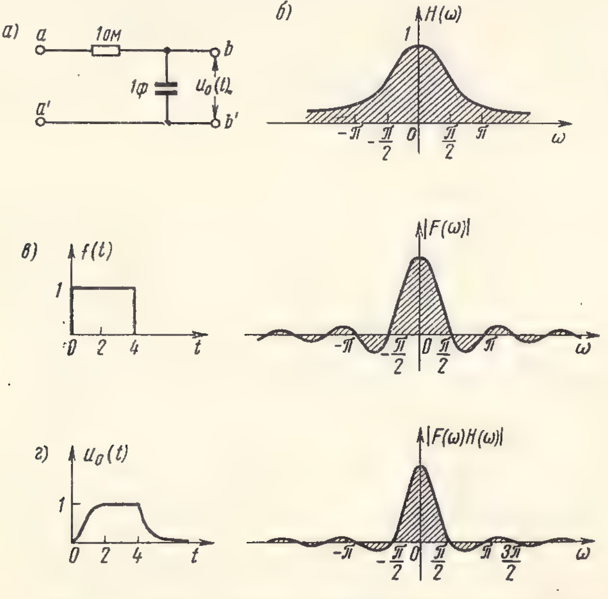


Рис. 2.2. Прохождение прямоугольного импульса через RC-цепь:

а) RC-цепь; *б)* модуль коэффициента передачи RC-цепи; *в)* входной сигнал и его спектр; *г)* сигнал на выходе RC-цепи и его спектр

плотности входного сигнала (прямоугольного импульса). Передаточная функция цепи $H(\omega)=1/(i\omega+1)$. График $|H(\omega)|$ приведен на рис. 2.2*б*. Фазовая характеристика цепи пока не рассматривается. Из приведенной характеристики видно, что цепь значитель-

но ослабляет высокочастотные составляющие сигнала и почти не ослабляет низкочастотные, поэтому она может служить простейшим фильтром нижних частот.

Модуль функции спектральной плотности отклика $|F(\omega)H(\omega)|$ показан на рис. 2.2*г*. Сравнение графиков 2.2*в* и 2.2*г* дает наглядное представление о затухании высокочастотных составляющих, вызываемом цепью. Очевидно, отклик $u_0(t)$ (рис. 2.2*г*) есть искаженная копия входного сигнала. Эти искажения вызваны тем, что цепь не одинаково пропускает все частотные составляющие входного сигнала[1]). Больше других ослабляются высокочастотные составляющие. Это проявляется в более медленном нарастании и спаде выходного напряжения по сравнению с входным. Входной сигнал скачком нарастает в момент $t=0$, что свидетельствует о наличии составляющих очень высоких частот. Так как цепь не пропускает высокочастотные составляющие, выходное напряжение не может изменяться с большой скоростью, в результате чего оно нарастает и спадает медленнее, чем входной сигнал.

2.3. Неискаженная передача

ХАРАКТЕРИСТИКИ НЕИСКАЖАЮЩЕЙ СИСТЕМЫ

Из изложенного вытекает, каким требованиям должна отвечать система, обеспечивающая неискаженную передачу сигналов. Прежде всего, система должна иметь постоянное затухание для всех частотных составляющих сигнала, т. е. передаточная функция $|H(\omega)|$ должна быть постоянна на всех частотах. Однако выполнение одного этого требования еще не гарантирует отсутствие искажений. Фазовый сдвиг каждой частотной составляющей также должен удовлетворять определенным соотношениям. Ранее влияние фазового сдвига не учитывалось. Если при передаче через систему частотные составляющие сигнала получают различные фазовые сдвиги, то даже при одинаковом затухании на всех частотах их сумма дает совершенно другой сигнал. Посмотрим, какому требованию должны удовлетворять относительные фазовые сдвиги составляющих, чтобы искажения отсутствовали.

Для неискаженной передачи необходимо, чтобы отклик был точной копией входного сигнала. Конечно, допускается различие в амплитуде, так как важна форма, а не величина отклика. Кроме того, выходной сигнал может запаздывать по времени относительно воздействия. Поэтому можно считать, что сигнал $f(t)$ передается без искажений, если отклик системы $r(t)=kf(t-t_0)$. На основании свойства (1.117) имеем

$$R(\omega)=F(\omega)H(\omega)=kF(\omega)e^{-i\omega t_0}.$$

---

[1]) Нарушая не только амплитудные, но и фазовые соотношения между отдельными частотными составляющими входного сигнала — *Прим. ред.*

Следовательно, неискажающая система должна иметь передаточную функцию

$$H(\omega)=k\,e^{-j\omega t_0}. \tag{2.6}$$

Эта передаточная функция показана на рис. 2.3*а*. Очевидно, амплитудно-частотная характеристика системы $|H(\omega)|$ постоянна на всех частотах и равна k. С другой стороны, фазовый сдвиг пропорционален частоте т. е.

$$\theta(\omega)=-\omega t_0. \tag{2.7а}$$

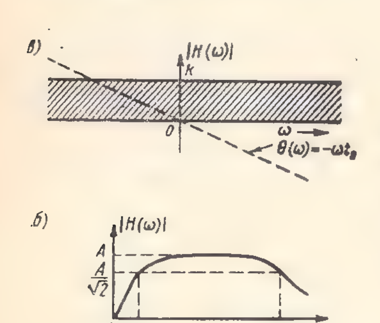


Рис. 2.3. Характеристики линейной системы:

а) идеальной (амплитудно-частотная и фазо-частотная); *б*) реальной (амплитудно-частотная характеристика)

Смысл этого очевиден: если две частотные составляющие сдвигаются на один и тот же временно́й интервал, их фазы изменяются пропорционально частоте. Например, если сигнал $\cos\omega t$ сдвинуть на t_0 *сек*, то получающийся сигнал $\cos\omega(t-t_0)$ можно записать как $\cos(\omega t-\omega t_0)$. Таким образом, новый сигнал имеет фазовый сдвиг $-\omega t_0$, пропорциональный ω.

Поскольку добавление фазового сдвига $n\pi$ может привести к изменению лишь знака сигнала, соотношение (2.7а) можно записать более строго:

$$\theta(\omega)=n\pi-\omega t_0, \tag{2.7б}$$

где n — целое число.

ПОЛОСА ПРОПУСКАНИЯ СИСТЕМЫ

Степень постоянства величины $|H(\omega)|$ системы обычно характеризуется ее полосой пропускания. Полоса пропускания системы произвольно определяется как интервал частот, в пределах которого $|H(\omega)|$ не становится меньше $1/\sqrt{2}$ своего значения на средней частоте[1]). Для системы, характеристика $|H(\omega)|$ которой показана на рис. 2.3*б*, полоса пропускания равна $\omega_2-\omega_1$.

Для неискаженной передачи система, по-видимому, должна обладать бесконечной полосой пропускания. В силу существующих физических ограничений такую систему создать невозможно. В действительности удовлетворительную неискаженную передачу можно получить в системе с ограниченной, но весьма большой полосой пропускания. Энергия любого физического сигнала убывает

---

[1]) Для простых цепей такое определение приводит к простым соотношениям, связывающим полосу пропускания с параметрами цепей.—*Прим. ред.*

с увеличением частоты, поэтому достаточно, чтобы система передавала лишь те частотные составляющие, в которых содержится наибольшая часть энергии сигнала. Ослабление высокочастотных составляющих должно привести к незначительным искажениям сигнала, так как эти составляющие несут весьма малую часть его энергии.

2.4. Идеальные фильтры

Идеальный фильтр нижних частот пропускает без искажений все сигналы, частоты которых не превышают определенную частоту *W, рад/сек*. Сигналы, имеющие частоты выше *W*, полностью по-

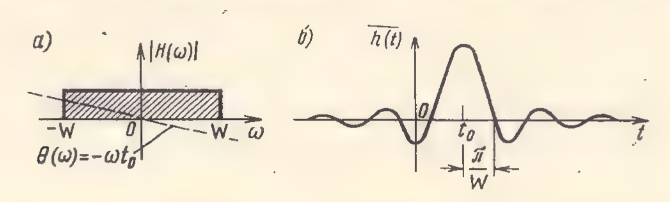


Рис. 2.4. Характеристики идеального фильтра нижних частот:
а) амплитудно-частотная и фазо-частотная; *б)* импульсная реакция

давляются (рис. 2.4*а*). Таким образом, амплитудно-частотная характеристика фильтра нижних частот представляет собой стробирующую функцию $G_{2W}(\omega)$. Соответствующая фазовая характеристика для неискаженной передачи равна $-\omega t_0$. Передаточная функция такого фильтра определяется соотношением

$$H(\omega) = |H(\omega)|\, e^{i\,\theta(\omega)} = G_{2W}(\omega)\, e^{-i\,\omega t_0}\,. \tag{2.8}$$

Реакцию такого фильтра $h(t)$ на единичный импульс можно найти как обратное преобразование Фурье $H(\omega)$:

$$h(t) = \mathscr{F}^{-1}[H(\omega)] = \mathscr{F}^{-1}\left[G_{2W}(\omega)\, e^{-i\,\omega t_0}\right].$$

Воспользовавшись парой преобразований *12* из табл. 1.1 и свойством временно́го сдвига, получим:

$$h(t) = \frac{W}{\pi} Sa\,[W(t - t_0)].$$

График этой функции на рис. 2.4*б* показывает, что импульсная реакция существует и при отрицательных значениях t. Это тем более непривычный результат, если учесть, что воздействие (единичный импульс) приложено в момент $t=0$. Отклик появляется раньше воздействия, которое система как бы предвосхищает. К сожалению, практически создать такую систему с предсказанием невоз-

можно. Следовательно, идеальный фильтр нижних частот физически нереализуем. Аналогичным образом можно показать, что и другие идеальные фильтры (такие, как идеальный фильтр верхних частот или идеальный полосовой фильтр, характеристики которых показаны на рис. 2.5) физически нереализуемы.

На практике удовлетворяются фильтрами, характеристики которых близки к идеальным. Простейший фильтр нижних частот показан на рис. 2.6*а*. Передаточная функция этого фильтра

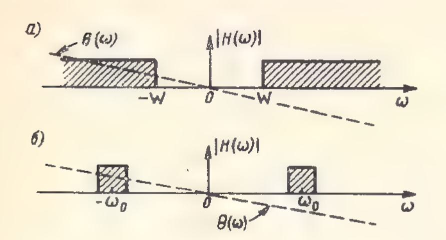
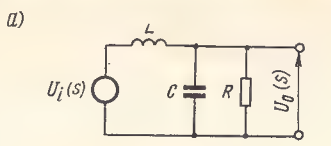
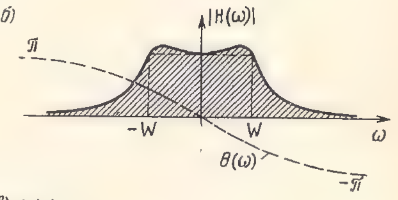


Рис. 2.5. Амплитудно- и фазо-частотные характеристики идеальных линейных систем:
а) фильтра верхних частот; *б)* полосового фильтра





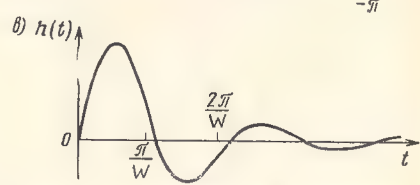


Рис. 2.6. Простейший реальный фильтр нижних частот:
а) схема; *б)* амплитудно- и фазо-частотные характеристики; *в)* импульсная реакция

$$H(\omega)=\frac{1/(1/R+i\omega C)}{i\omega L+1/(1/R+i\omega C)}=\frac{1}{1-\omega^2 LC+i\omega L/R}.$$

Так как $1/\sqrt{LC}=W$ и $\sqrt{L/C}=R$,

$$H(\omega)=\frac{\frac{\sqrt{3}}{2}W}{\left(\frac{W}{2}+i\omega\right)^2+\left(\frac{\sqrt{3}}{2}W\right)^2}\frac{2W}{\sqrt{3}}.$$

Импульсная реакция $h(t)$ имеет вид (пара преобразований *11*, табл. 1.1):

$$h(t)=\mathcal{F}^{-1}[H(\omega)]=\frac{2W}{\sqrt{3}}e^{-\frac{Wt}{2}}\sin\left(\frac{\sqrt{3}}{2}Wt\right).$$

На рис. 2.6*б* показаны амплитудно- и фазо-частотные характеристики этого фильтра, а на рис. 2.6*в* — импульсная реакция $h(t)$. Импульсная реакция близка к реакции идеального фильтра, только она начинается в момент $t=0$.

Желательно найти признак, по которому можно было бы различать физически реализуемые и физически нереализуемые характеристики. Таким признаком является критерий Пэйли-Винера.

2.5. Причинность и физическая реализуемость. Критерий Пэйли-Винера

В литературе физическая реализуемость определяется различными способами. Воспользуемся наиболее общим определением, позволяющим различать два отмеченных класса систем. Физически реализуемая система не может порождать отклик, начинающийся раньше момента приложения воздействия. Приведенный довод известен под названием условия причинности. Сформулировать это условие можно двояким образом. Отклик $h(t)$ на единичный импульс физически реализуемой системы должен быть равен нулю при $t<0$. Это временно́й критерий физической реализуемости. В частотной области указанный критерий формулируется следующим образом: для физической реализуемости системы необходимо и достаточно выполнение условия [1])

$$\int_{-\infty}^{\infty} \frac{|\ln H(\omega)|}{1+\omega^2} d\omega < \infty. \quad (2.9)$$

Однако для выполнения критерия Пэйли-Винера амплитудно-частотная характеристика $H(\omega)$ должна быть интегрируемой в квадрате [2]), т. е. $\int_{-\infty}^{\infty} |H(\omega)|^2 d\omega < \infty$. Если амплитудно-частотная характеристика не удовлетворяет критерию Пэйли-Винера (2.9), то система имеет непричинную импульсную реакцию, т. е. реакцию, существующую до того, как к системе приложено воздействие.

Из ф-лы (2.9) следует, что амплитудно-частотная характеристика $|H(\omega)|$ может быть равной нулю на некоторых дискретных частотах, но не может оставаться равной нулю в конечной полосе частот, так как это приведет к расходящемуся интегралу. Отсюда можно заключить, что идеальные фильтры, характеристики которых приведены на рис. 2.4 и 2.5, физически нереализуемы. Из ф-лы (2.9) вытекает, что амплитудно-частотная характеристика не может уменьшаться до нуля быстрее, чем экспонента. Итак, функция $|H(\omega)| = ke^{-\alpha|\omega|}$ допустима, а гауссова функция $|H(\omega)| = ke^{-\alpha\omega^2}$ относится к физически нереализуемой системе, так как

---

[1]) Raymond E. A., Paley and N. Wiener. Fourier Transform in the Complex domain. Am Mathemat. Soc. Colloquium Publication 19, New. York, 1934.

[2]) Если амплитудно-частотная характеристика $|H(\omega)|$ удовлетворяет критерию Пэйли-Винера (2.9), то из этого еще не следует, что система физически реализуема. Необходимо выбрать также соответствующую фазо-частотную характеристику, чтобы результирующая функция была физически реализуемой.

условие (2.9) не удовлетворяется. Интересно отметить, что хотя идеальные характеристики, показанные на рис. 2.4 и 2.5, нереализуемы, практически можно получить характеристики, сколь угодно близкие к идеальным. Так, характеристика фильтра нижних частот, показанная на рис. 2.7, физически реализуема при сколь угодно малых величинах ε. Нетрудно убедиться в том, что эта характеристика удовлетворяет критерию Пэйли-Винера.

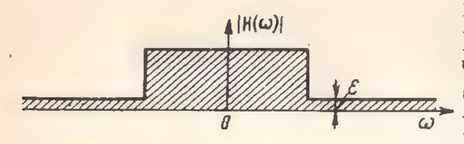


Рис. 2.7. Характеристика физически реализуемого фильтра

2.6. Связь между полосой пропускания и временем установления

Быстрые изменения мгновенных значений сигнала во времени вызывают появление в спектре сигнала высокочастотных составляющих. Сигнал, изменяющийся сравнительно медленно, содержит преимущественно низкочастотные составляющие. Выше было показано (рис. 2.2), что если через фильтр нижних частот проходит сигнал с разрывом непрерывности, то в выходном сигнале разрывность сглаживается. При подаче на вход идеального фильтра нижних частот единичного скачка $u_1(t)$ выходное напряжение нарастает постепенно (а не скачкообразно, как на входе). Время нарастания выходного напряжения зависит от частоты среза фильтра. Покажем, что время нарастания выходного напряжения обратно пропорционально частоте среза фильтра: чем меньше частота среза, тем более плавно нарастает выходной сигнал.

В этом легко убедиться, рассмотрев реакцию идеального фильтра нижних частот на единичный скачок $u_1(t)$. Для идеального фильтра нижних частот передаточная функция (рис. 2.4*а*) определяется выражением (2.8). Кроме того,

$$u_1(t) \leftarrow\rightarrow \pi\delta(\omega) + \frac{1}{\mathrm{i}\,\omega}.$$

Если $r(t)$ — реакция фильтра на $u_1(t)$, то $R(\omega) = \left[\pi\delta(\omega) + \frac{1}{\mathrm{i}\,\omega}\right] G_{2W}(\omega)\, \mathrm{e}^{-\mathrm{i}\,\omega t_0}$.

Заметим, что $H(0) = 1$. Следовательно, $\delta(\omega) H(\omega) = H(0)\,\delta(\omega)$ и $R(\omega) = \pi\delta(\omega) + \frac{1}{\mathrm{i}\,\omega} G_{2W}(\omega)\, \mathrm{e}^{-\mathrm{i}\,\omega t_0}$,

$$r(t) = \mathcal{F}^{-1}\left[\pi\delta(\omega) + \frac{1}{\mathrm{i}\,\omega} G_{2W}(\omega)\, \mathrm{e}^{-\mathrm{i}\,\omega t_0}\right] = \frac{1}{2} + \mathcal{F}^{-1}\left[\frac{1}{\mathrm{i}\,\omega} G_{2W}(\omega)\, \mathrm{e}^{-\mathrm{i}\,\omega t_0}\right] =$$

$$= \frac{1}{2} + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{G_{2W}(\omega)}{\mathrm{i}\,\omega}\, \mathrm{e}^{\mathrm{i}\,\omega(t-t_0)}\, d\omega = \frac{1}{2} + \frac{1}{2\pi}\int_{-W}^{W} \frac{\mathrm{e}^{\mathrm{i}\,\omega(t-t_0)}}{\mathrm{i}\,\omega}\, d\omega =$$

$$= \frac{1}{2} + \frac{1}{2\pi} \int_{-W}^{W} \frac{\cos [\omega (t - t_0)]}{i\omega} d\omega + \frac{1}{2\pi} \int_{-W}^{W} \frac{\sin [\omega (t - t_0)]}{\omega} d\omega.$$

Первый интеграл этого выражения равен нулю, так как подынтегральное выражение — нечетная функция ω. Подынтегральная функция во втором интеграле — четная, поэтому

$$r(t) = \frac{1}{2} + \frac{1}{\pi} \int_{0}^{W} \frac{\sin [\omega (t - t_0)]}{\omega} d\omega =$$

$$= \frac{1}{2} + \frac{1}{\pi} \int_{0}^{W(t-t_0)} \frac{\sin x}{x} dx = \frac{1}{2} + \frac{1}{\pi} \int_{0}^{W(t-t_0)} Sa(x)\, dx.$$

Рис. 2.8. Функция $\mathrm{Si}(x)$:

а) график функции $\mathrm{Si}(t)$; *б)* функция единичного скачка напряжения; *в)* отклик идеального фильтра нижних частот на единичный скачок напряжения

Последний интеграл в замкнутой форме не выражается. Он табулирован[1]) и называется интегральным синусом

$$\mathrm{Si}(x) = \int_{0}^{x} Sa(\tau)\, d\tau. \tag{2.10а}$$

График функции $\mathrm{Si}(x)$ приведен на рис. 2.8*а*. Теперь реакция $r(t)$ записывается в виде:

$$r(t) = \frac{1}{2} + \frac{1}{\pi} \mathrm{Si}[W(t - t_0)]. \tag{2.10б}$$

На рис. 2.8*б, в* показаны соответственно единичный скачок $u_1(t)$ и реакция на него $r(t)$. Частота среза фильтра нижних частот равна W (рис. 2.4*а*). Из рис. 2.8*в* видно, что с уменьшением частоты среза выходное напряжение $r(t)$ нарастает более медленно. Если

[1]) Е. Янке и Ф. Эмде. Таблицы функций с формулами и кривыми. ГИТТЛ, 1949.

определить время нарастания t_r как время, в течение которого выходное напряжение изменяется от минимального до максимального значения [1]), то

$$t_r = \frac{2\pi}{W} = \frac{1}{B}, \qquad (2.11)$$

где B — частота среза фильтра в герцах.

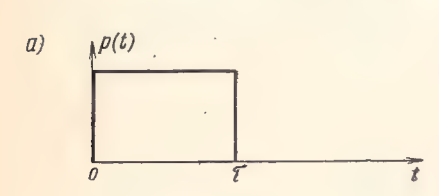
На основании полученных выводов легко найти реакцию фильтра нижних частот на прямоугольный импульс $p(t)$, изображенный на рис. 2.9*а*;

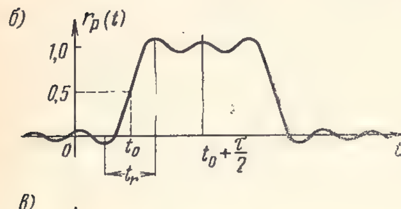
$$p(t) = u_1(t) - u_1(t-\tau).$$

Реакцию $r_p(t)$ фильтра на этот импульс можно получить из соотношения (2.10б), используя принцип суперпозиции:

$$r_p(t) = \frac{1}{\pi}\{\mathrm{Si}[W(t-t_0)] - \\ - \mathrm{Si}[W(t-t_0-\tau)]\}.$$

График выходного напряжения показан на рис. 2.9*б*. На рис. 2.9*в* приведена приемлемая аппроксимация этого напряжения трапецеидальным импульсом. Из рисунка видно, что передача импульса через фильтр нижних частот вызывает его дисперсию (расплывание) во времени.





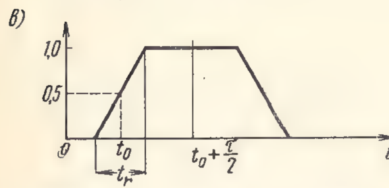


Рис. 2.9 Отклик фильтра нижних частот на прямоугольный импульс: *а*) воздействие; *б*) отклик; *в*) возможная аппроксимация отклика

2.7. Спектр плотности энергии

ЭНЕРГИЯ СИГНАЛА

Удобным параметром сигнала $f(t)$ является его нормализованная энергия (или просто энергия) E, которая определяется как энергия, рассеиваемая напряжением $f(t)$, приложенным к сопро-

[1]) В литературе существуют различные определения времени нарастания. Иногда оно определяется как время, в течение которого реакция нарастает от нулевого (ближайшего к минимальному) значения до максимального. Тогда $t_r = 0{,}8/B$. Существует определение времени нарастания как величины, обратной наклону $r(t)$ в момент $t = t_0$. В этом случае $t_r = 0{,}5/B$. В импульсной технике t_r определяется как время, в течение которого процесс нарастает от 0,1 до 0,9 установившегося значения; при этом $t_r \approx 0{,}44/B$. Таким образом, во всех определениях t_r обратно пропорционально частоте среза.

тивлению в 1 *ом* (или током $f(t)$, протекающим через сопротивление в 1 *ом*):

$$E=\int_{-\infty}^{\infty} f^2(t)\,dt. \qquad (2.12)$$

Понятие энергии сигнала имеет смысл только в том случае, если конечен интеграл (2.12). Сигналы с конечной энергией называются энергетическими (а также импульсными). Для некоторых сигналов (например, периодических) интеграл (2.12) равен бесконечности и понятие энергии для них теряет смысл. В этом случае рассматривают среднюю по времени энергию, т. е. среднюю мощность сигнала. Сигналы с ограниченной мощностью, называемые кратко мощностными, будут рассматриваться ниже.

Если $F(\omega)$ — преобразование Фурье сигнала $f(t)$, то используя ф-лу (1.73), можно записать

$$E=\int_{-\infty}^{\infty} f^2(t)\,dt=\int_{-\infty}^{\infty} f(t)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\,e^{i\omega t}\,d\omega\right]dt.$$

Изменив порядок интегрирования в правой части этого выражения, получим

$$E=\int_{-\infty}^{\infty} f^2(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\left[\int_{-\infty}^{\infty} f(t)\,e^{i\omega t}\,dt\right]d\omega.$$

Внутренний интеграл в правой части полученного выражения равен $F(-\omega)$, поэтому

$$\int_{-\infty}^{\infty} f^2(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\,F(-\omega)\,d\omega.$$

Для действительной функции $f(t)$ на основании ф-лы (1.77) имеем $F(\omega)F(-\omega)=|F(\omega)|^2$ и

$$\int_{-\infty}^{\infty} f^2(t)\,dt=\frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\,d\omega=\int_{-\infty}^{\infty} |F(\omega)|^2\,df. \qquad (2.13)$$

Полученное соотношение[1]) устанавливает, что энергия сигнала равна площади под кривой $|F(\omega)|^2$ (интегрирование ведется по переменной $f=\omega/2\pi$).

---

[1]) Выражение (2.13) называется теоремой Парсеваля или теоремой Планшерля и является аналогом ф-лы (1.34) для непериодических функций.

Плотность энергии имеет весьма интересное физическое толкование. Рассмотрим сигнал $f(t)$, поступающий на вход идеального полосового фильтра с передаточной функцией $H(\omega)$, показанной на рис. 2.10. Этот фильтр подавляет все частотные составляющие, за исключением узкой полосы $\Delta\omega (\Delta\omega \to 0)$, вблизи ω_0. Преобразование Фурье отклика $r(t)$ фильтра $R(\omega) = F(\omega) H(\omega)$, энергия E_0 выходного сигнала $r(t)$

$$E_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega) H(\omega)|^2 d\omega.$$

Так как функция $H(\omega) = 0$ везде, кроме узкой полосы $\Delta\omega$, где она равна единице, можно записать (при $\Delta\omega \to 0$):

$$F_0 = 2 \frac{1}{2\pi} |F(\omega)|^2 \Delta\omega.$$

Таким образом, энергия выходного сигнала равна $2|F(\omega)|^2 \Delta f$. Так как все частотные составляющие сигнала, кроме попадающих в узкую полосу Δf, подавляются фильтром, то $2|F(\omega)|^2$ есть энергия сигнала, приходящаяся на единицу полосы Δf вблизи ω_0. Заметим, что единицей спектральной плотности энергии служит *дж/гц*. Следует также отметить, что энергия сосредоточена в составляющих как с положительными, так и с отрицательными частотами. Более того, энергии на положительных и на отрицательных частотах равны, так как $|F(\omega)|^2 = |F(-\omega)|^2$.

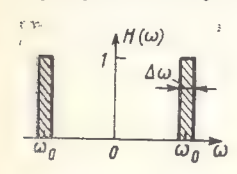


Рис. 2.10. Функция передачи идеального полосового фильтра

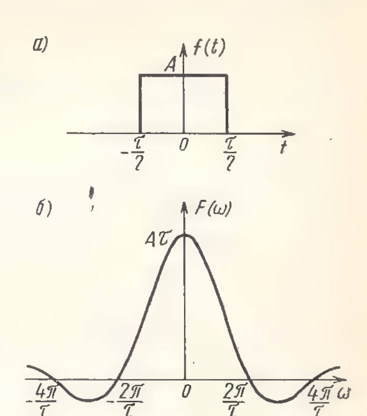
Таким образом, можно считать, что $|F(\omega)|^2$ (половина энергии $2|F(\omega)|^2$) является энергией составляющих на положительных частотах, а оставшаяся половина $|F(\omega)|^2$ — их энергией на отрицательных частотах[1]). Поэтому $|F(\omega)|^2$ называют спектром плотности энергии, т. е. энергией, приходящейся на единицу полосы (положительной или отрицательной). Таким образом, спектр плотности энергии определяется как функция[2])

$$\Psi_f(\omega) = |F(\omega)|^2. \tag{2.14}$$

[1]) Разделение полной энергии на энергию положительных и отрицательных частот вводится для удобства. В действительности сумма энергий составляющих на положительных и отрицательных частотах дает полную энергию любой части частотного диапазона.

[2]) В литературе $\Psi_f(\omega)$ определяют различными способами. Например, $\Psi_f(\omega) = 2|F(\omega)|^2$, т. е. энергии на положительных и отрицательных частотах суммируются. Полная энергия E находится интегрированием $\Psi_f(\omega)$ в пределах от 0 до ∞ или от 0 до $-\infty$. Согласно другому определению $\Psi_f(\omega) = \frac{1}{\pi}|F(\omega)|^2$. В этом случае плотность энергии определяется как энергия, приходящаяся на полосу в 1 *рад/сек*.

Спектр плотности энергии показывает относительный вклад различных частотных составляющих в общую энергию. На рис. 2.11 показаны прямоугольная функция, ее преобразование Фурье и спектр плотности энергии $|F(\omega)|^2$. Полная энергия

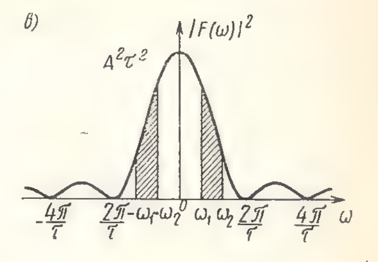


$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega =$$

$$= \int_{-\infty}^{\infty} |F(\omega)|^2 df = \int_{-\infty}^{\infty} \Psi_f(\omega)\, df. \qquad (2.15)$$

Так как $|F(\omega)|^2 = |F(-\omega)|^2$, спектр плотности энергии — действительная и четная функция частоты. Следовательно, выражение (2.15) можно записать в виде

$$E = \frac{1}{\pi}\int_0^{\infty} F(\omega)^2 d\omega = 2\int_0^{\infty} |F(\omega)|^2 df. \qquad (2.16)$$



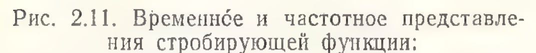


Рис. 2.11. Временнóе и частотное представления стробирующей функции:
а) стробирующая функция; *б)* функция спектральной плотности; *в)* спектр плотности энергии

ПЛОТНОСТЬ ЭНЕРГИИ ВЫХОДНОГО СИГНАЛА

Если $f(t)$ и $r(t)$ — соответственно воздействие и отклик линейной системы с передаточной функцией $H(\omega)$, то $R(\omega) = H(\omega)F(\omega)$. Так как $|F(\omega)|^2$ является спектром плотности энергии воздействия, а $|R(\omega)|^2$ — аналогичная функция отклика,

$$|R(\omega)|^2 = |H(\omega)|^2\, |F(\omega)|^2 = |H(\omega)|^2\, \Psi_f(\omega). \qquad (2.17)$$

2.8. Спектр плотности мощности

ОПРЕДЕЛЕНИЕ СПЕКТРА ПЛОТНОСТИ МОЩНОСТИ

Выше отмечалось, что некоторые сигналы (например, периодические) имеют бесконечную энергию. Такие сигналы называются мощностными. Определяющим параметром мощностного сигнала $f(t)$ является средняя мощность P. Определим среднюю мощность (или просто мощность) сигнала $f(t)$ как среднюю мощность, рассеиваемую напряжением $f(t)$ на сопротивлении в 1 *ом* (или током, протекающим через сопротивление в 1 *ом*):

$$P=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} f^2(t)\,dt. \qquad (2.18)$$

Заметим, что средняя мощность, определяемая ф-лой (2.18), есть среднее значение квадрата $f(t)$, обозначаемое как $\overline{f^2(t)}$:

$$P=\overline{f^2(t)}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} f^2(t)\,dt. \qquad (2.19)$$

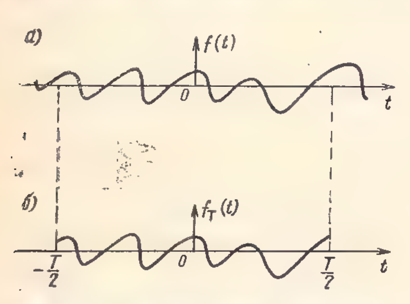


Рис. 2.12. Мощностные сигналы: *а)* неограниченный во времени сигнал; *б)* сигнал конечной длительности

Дальнейший анализ аналогичен анализу плотности энергии. Образуем новую функцию $f_T(t)$, ограничив $f(t)$ интервалом $|t|<T/2$. Такую ограниченную функцию (рис. 2.12*б*) можно записать в виде

$$f_T(t)=\begin{cases} f(t) & \text{при } |t|<T/2, \\ 0 & \text{при } |t|>T/2. \end{cases}$$

При конечном T функция $f_T(t)$ имеет ограниченную энергию. Пусть $f_T(t)\longleftrightarrow F_T(\omega)$. Тогда энергия E_T сигнала $f_T(t)$:

$$E_T=\int_{-T/2}^{T/2} f_T^2(t)\,dt=\int_{-\infty}^{\infty}|F_T(\omega)|^2 df,$$

поэтому средняя мощность

$$P=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} f^2(t)\,dt=\int_{-\infty}^{\infty}\lim_{T\to\infty}\frac{|F_T(\omega)|^2}{T}\,df.$$

С увеличением T энергия $f_T(t)$ и величина $|F_T(\omega)|^2$ возрастают, однако отношение $|F_T(\omega)|^2/T$ может стремиться к пределу. Предположим, что такой предел существует, и найдем спектр плотности мощности $S_f(\omega)$ сигнала $f(t)$ как

$$S_f(\omega)=\lim_{T\to\infty}|F_T(\omega)|^2/T. \qquad (2.20)$$

Следовательно,

$$P=\overline{f^2(t)}=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} f^2(t)\,dt=\int_{-\infty}^{\infty}S_f(\omega)df=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_f(\omega)\,d\omega. \qquad (2.21)$$

Из ф-лы (2.20) следует, что спектр плотности мощности является четной функцией ω, поэтому ф-лу (2.21) можно записать в виде

$$P=\overline{f^2(t)}=2\int\limits_0^\infty S_f(\omega)\,df=\frac{1}{\pi}\int\limits_0^\infty S_f(\omega)\,d\omega. \qquad (2.22)$$

Спектр плотности мощности сигнала сохраняет информацию только об амплитудах спектральных составляющих $F_T(\omega)$, информация о фазе теряется. Отсюда следует, что все сигналы с одинаковым спектром амплитуд и различными спектрами фаз будут иметь одинаковые спектры плотности мощности.

Итак, данному сигналу соответствует единственный спектр плотности мощности. Обратное утверждение не верно; может существовать большое число сигналов (фактически бесконечное), имеющих один и тот же спектр плотности мощности.

Рассмотрим, например, сигналы $f(t)\cos\omega_0 t$ и $f(t)\sin\omega_0 t$, спектры которых для случая, когда $f(t)$ низкочастотный энергетический сигнал, выражаются ф-лами (1.116а) и (1.116б). Если теперь $f(t)$ — мощностной сигнал со спектром плотности мощности $S_f(\omega)$, то спектр плотности мощности $\varphi(t)=f(t)\cos\omega_0 t$ по определению

$$S_\varphi(\omega)=\lim_{T\to\infty}|\Phi_T(\omega)|^2/T,$$

где

$$f_T(t)\cos\omega_0 t=\varphi_T(t)\longleftrightarrow\Phi_T(\omega).$$

На основании (1.116а) имеем:

$$\Phi_T(\omega)=\frac{1}{2}[F_T(\omega+\omega_0)+F_T(\omega-\omega_0)],$$

$$S_\varphi(\omega)=\frac{1}{4}\lim_{T\to\infty}\frac{|F_T(\omega+\omega_0)+F_T(\omega-\omega_0)|^2}{T}=$$

$$=\frac{1}{4}\lim_{T\to\infty}\frac{|F_T(\omega+\omega_0)+F_T(\omega-\omega_0)|^2}{T}=$$

$$=\frac{1}{4}\lim_{T\to\infty}\frac{|F_T(\omega+\omega_0)|^2+|F_T(\omega-\omega_0)|^2}{T}. \qquad (2.23)$$

Заметим, что произведение $F_T(\omega+\omega_0)F_T(\omega-\omega_0)$ равно нулю, так как эти два спектра не перекрываются (рис. 2.13). Поэтому спектр сигнала $\varphi(t)$ определяется выражением[1])

---

[1]) Полученные здесь результаты применимы к низкочастотным сигналам $f(t)$, для которых спектры $F_T(\omega+\omega_0)$ и $F_T(\omega-\omega_0)$ не перекрываются. Если $f(t)$ — сигнал с полосовым спектром, то $F_T(\omega+\omega_0)$ и $F_T(\omega-\omega_0)$ могут частично перекрываться вблизи начала координат и, строго говоря, ф-ла (2.24) будет неприменима. Однако если $f(t)$ — случайный сигнал, то эту формулу применять можно.

Отдельные реализации случайного процесса представляют собой колебания различной формы, поэтому спектр плотности мощности случайного сигнала (случайного процесса) определяется как среднее значение спектров всех возможных реализаций процесса. В этом случае можно показать, что хотя $F_T(\omega+\omega_0)$ и $F_T(\omega-\omega_0)$ перекрываются, среднее значение их произведения равно нулю и ф-ла (2.24) справедлива.

$$S_\varphi(\omega) = \frac{1}{4}\left[S_f(\omega + \omega_0) + S_f(\omega - \omega_0)\right]. \qquad (2.24)$$

Аналогичный результат получается и для сигнала $\varphi(t) = f(t)\sin\omega_0 t$.

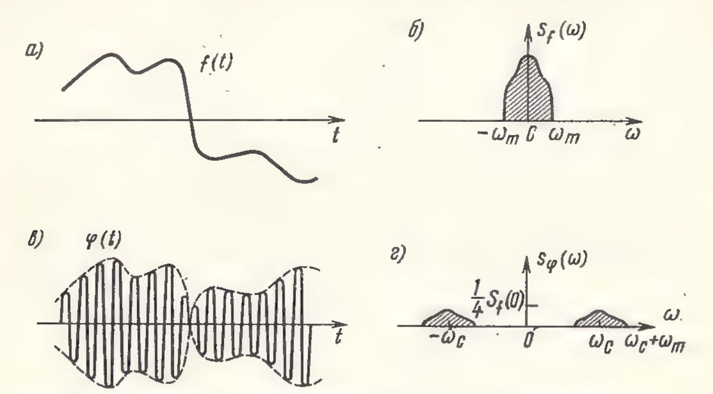


Рис. 2.13. К определению мощности амплитудномодулированного сигнала:
а) модулирующая функция; *б)* спектр плотности мощности модулирующей функции; *в)* амплитудномодулированный сигнал; *г)* спектр плотности мощности модулированного сигнала

Итак, сигналы $f(t)\cos\omega_0 t$ и $f(t)\sin\omega_0 t$ имеют одинаковые спектры плотности мощности. На рис. 2.13 показаны спектры плотности мощности сигналов $f(t)$ и $\varphi(t)$. Формула (2.24) есть не что иное, как теорема о модуляции применительно к мощностным сигналам.

Пример 2.1.

Вычислим мощность (среднее квадратическое значение) амплитудномодулированного сигнала $\varphi(t) = f(t)\cos\omega_c t$, где $f(t)$ — мощностной сигнал со средним квадратическим значением $\overline{f^2(t)}$.

Мощность сигнала равна площади под кривой спектра плотности мощности, деленной на 2π:

$$\overline{\varphi^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_\varphi(\omega)\, d\omega, \quad \overline{f^2(t)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_f(\omega)\, d\omega.$$

Площадь под кривой $S_f(\omega)$ равна $2\pi\overline{f^2(t)}$ (рис. 2.13). Но из ф-лы (2.24) следует, что площадь под кривой $S_\varphi(\omega)$ равна половине площади под кривой $S_f(\omega)$. Это также видно из рис. 2.13. Следовательно.

$$\overline{\varphi^2(t)} = \overline{[f(t)\cos\omega_c t]^2} = \frac{1}{2}\overline{f^2(t)}. \qquad (2.25)$$

Аналогичный результат получается для $\varphi(t) = f(t)\sin\omega_c t$.

При выводе ф-лы (2.25) предполагалось, что $\omega_c > \omega_m$. Если $\omega_c < \omega_m$, то спектры $S_f(\omega + \omega_c)$ и $S_f(\omega - \omega_c)$ на рис. 2.13*г* будут перекрываться и соотношение (2.24) нарушится. Таким образом, ф-ла (2.25) верна только в том случае, если ω_c больше наивысшей частоты в спектре $f(t)$.

Пример 2.2.

Выразим среднюю мощность сигнала с ограниченным спектром через его отсчеты. Сигнал $f(t)$, спектр которого ограничен частотой ω_m по ф-ле (1.140), представляется как

$$f(t)=\sum_k f_k\, Sa\,[\omega_m(t-kT)], \tag{2.26}$$

где $f_k=f(kT)$ — k-й отсчет, T — интервал между отсчетами

$$T=1/2f_m=\pi/\omega_m. \tag{2.27}$$

Средняя мощность сигнала $f(t)$

$$\overline{f^2(t)}=\lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_{-\tau/2}^{\tau/2} f^2(t)\,dt.$$

Подставляя сюда ф-лу (2.26) и меняя порядок суммирования и интегрирования, получаем

$$\overline{f^2(t)}=\lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_{-\tau/2}^{\tau/2}\left(\sum_k f_k\, Sa\,[\omega_m(t-kT)]\right)^2 dt. \tag{2.28}$$

Можно показать что различные функции отсчетов ортогональны, т. е.

$$\int\limits_{-\infty}^{\infty} Sa\,[\omega_m(t-nT)]\, Sa\,[\omega_m(t-mT)]\,dt=\begin{cases}\dfrac{\pi}{\omega_m} & \text{при} \quad m=n,\\ 0 & \text{при} \quad m\neq n.\end{cases} \tag{2.29}$$

Следовательно,

$$\overline{f^2(t)}=\sum_k \lim_{\tau\to\infty}\frac{1}{\tau}\int\limits_{-\tau/2}^{\tau/2} f_k^2\, Sa^2\,[\omega_m(t-kT)]\,dt=$$

$$=\sum_k \frac{f_k^2}{\tau}\lim_{\tau\to\infty}\int\limits_{-\tau/2}^{\tau/2} Sa^2\,[\omega_m(t-kT)]\,dt=\sum_k\frac{\pi}{\omega_m\tau}f_k^2=\frac{1}{2f_m\tau}\sum_k f_k^2. \tag{2.30}$$

Заметим что τ — интервал, по которому выполняется усреднение. Поскольку частота отсчетов равна $2f_m$ отсчетов в секунду, то $2f_m\tau$ есть общее число отсчетов на интервале τ. Правая часть ф-лы (2.30) представляет среднее значение квадрата отсчетов, т. е. среднее значение квадрата мгновенных значений сигнала. Следовательно, средняя мощность сигнала с ограниченным спектром равна среднему значению квадрата его отсчетов:

$$\overline{f^2(t)}=\overline{f_k^2}. \tag{2.31}$$

СПЕКТР ПЛОТНОСТИ МОЩНОСТИ ПЕРИОДИЧЕСКОГО СИГНАЛА

Преобразование Фу ье $F(\omega)$ периодической функции $f(t)$ определяется по ф-ле (1.102): $F(\omega)=2\pi\sum\limits_{n=-\infty}^{\infty}F_n\delta(\omega-n\omega_0)$. Для ограни-

ченной по времени функции $f_T(t)=G_T(t)f(t)$ на основании теоремы частотной свертки имеем

$$F_T(\omega)=\frac{1}{2\pi}TSa\left(\frac{\omega T}{2}\right)*F(\omega)=TSa\left(\frac{\omega T}{2}\right)*\sum_{n=-\infty}^{\infty}F_n\delta(\omega-n\omega_0)=$$

$$=T\sum_{n=-\infty}^{\infty}F_nSa\left(\frac{\omega T}{2}\right)*\delta(\omega-n\omega_0)=T\sum_{n=-\infty}^{\infty}F_nSa\left[\frac{(\omega-n\omega_0)T}{2}\right].$$

Поэтому

$$\lim_{T\to\infty}\frac{1}{T}|F_T(\omega)|^2=\lim_{T\to\infty}T\sum_{n=-\infty}^{\infty}|F_n|^2 Sa^2\left[\frac{(\omega-n\omega_0)T}{2}\right]. \qquad (2.32)$$

Заметим, что при $T\to\infty$ функция $Sa\{[(\omega-n\omega_0)T]/2\}$ концентрируется возле $\omega=n\omega_0$ и в ф-ле (2.32) не содержатся слагаемые взаимных произведений, поскольку каждый из сомножителей существует на частотах, где все остальные сомножители равны нулю. Воспользовавшись ф-лой (1.86), получим

$$S_f(\omega)=\lim_{T\to\infty}\frac{|F_T(\omega)|^2}{T}=2\pi\sum_{n=-\infty}^{\infty}|F_n|^2\delta(\omega-n\omega_0). \qquad (2.33)$$

Следовательно, если

$$f(t)=a\cos(\omega_0t+\theta)=\frac{a}{2}\left[e^{i(\omega_0t+\theta)}+e^{-i(\omega_0t+\theta)}\right]=$$

$$=\left(\frac{a}{2}e^{i\theta}\right)e^{i\omega_0t}+\left(\frac{a}{2}e^{-i\theta}\right)e^{-i\omega_0t},$$

то

$$S_f(\omega)=2\pi\left[\frac{a^2}{4}\delta(\omega-\omega_0)+\frac{a^2}{4}\delta(\omega+\omega_0)\right]=$$

$$=\frac{\pi a^2}{2}[\delta(\omega-\omega_0)+\delta(\omega+\omega_0)]. \qquad (2.34)$$

Таким образом, спектр плотности мощности сигнала $a\cos(\omega_0t+\theta)$ состоит из двух единичных импульсов на частотах $\pm\omega_0$ интенсивностью $\pi a^2/2$ каждый. Заметим, что спектр плотности мощности не зависит от θ.

Мощность рассматриваемого сигнала

$$P=\overline{f^2(t)}=a^2/2, \qquad (2.35)$$

так как среднее значение квадрата любого синусоидального сигнала с амплитудой a равно $a^2/2$. С другой стороны, мощность P равна площади под кривой $S_f(\omega)$, деленной на 2π, что также дает $a^2/2$.

Пример 2.3

Найдем мощность сигнала $A+f(t)$, где A — постоянная, а $f(t)$ — мощностной сигнал с нулевым средним значением:

$$P=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}[A+f(t)]^2\,dt=\lim_{T\to\infty}\frac{1}{T}\left[\int_{-T/2}^{T/2}A^2\,dt+\int_{-T/2}^{T/2}f^2(t)\,dt+\right.$$

$$\left.+2A\int_{-T/2}^{T/2}f(t)\,dt\right]=A^2+\overline{f^2(t)}+\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}f(t)\,dt.$$

Интеграл в правой части этого выражения представляет собой среднее значение $f(t)$, равное по условию нулю. Следовательно,

$$P=\overline{[A+f(t)]^2}=A^2+\overline{f^2(t)}. \qquad (2.36)$$

СПЕКТРЫ ПЛОТНОСТИ МОЩНОСТИ ВОЗДЕЙСТВИЯ И ОТКЛИКА

Пусть на вход линейной системы с передаточной функцией $H(\omega)$ поступает мощностной сигнал $f(t)$. Выходной сигнал обозначим через $r(t)$, записав связь между входом и выходом символически $f(t)\to r(t)$. Обозначим через $f_T(t)$ и $r_T(t)$ сигналы, совпадающие с $f(t)$ и $r(t)$ на интервале времени $|t|=T/2$ и равные нулю вне его.

Допустим, что на вход системы поступает сигнал $f_T(t)$. Это равносильно тому. что на вход системы поступает сигнал $f(t)$ в течение времени $|t|\leqslant T/2$, а в остальное время этот сигнал равен нулю. Отклик в общем случае не равен $r_T(t)$; он существует и при $t>T/2$. Однако, поскольку входной сигнал равен нулю при $t>T/2$, отклик при $t>T/2$ должен затухать во времени. В пределе при $T\to\infty$ частью функции вне $t=T/2$ можно пренебречь[1]). Следовательно, при $T\to\infty$ без большой погрешности можно считать, что откликом на воздействие $f_T(t)$ является $r_T(t)$:

$$\lim_{T\to\infty}f_T(t)\to r_T(t),\quad \lim_{T\to\infty}R_T(\omega)=H(\omega)F_T(\omega).$$

Кроме того, по определению спектр плотности мощности выходного сигнала $r(t)$

$$S_r(\omega)=\lim_{T\to\infty}\frac{1}{T}|R_T(\omega)|^2=\lim_{T\to\infty}\frac{1}{T}|H(\omega)F_T(\omega)|^2=$$

$$=|H(\omega)|^2\lim_{R\to\infty}\frac{1}{T}|F_T(\omega)|^2=|H(\omega)|^2S_f(\omega). \qquad (2.37)$$

---

[1]) Сходное доказательство применяется и для случая $t<-T/2$. В этом случае отклик, вызванный действием сигнала $f(t)$, при $t<-T/2$ становится несущественным, если $T\to\infty$.

Среднее значение квадрата сигнала равно деленной на 2π площади, ограниченной спектром плотности мощности. Следовательно, средняя мощность отклика $r(t)$

$$\overline{r^2(t)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 S_f(\omega)\, d\,\omega.$$

ФИЗИЧЕСКОЕ ТОЛКОВАНИЕ СПЕКТРА ПЛОТНОСТИ МОЩНОСТИ

Спектр плотности мощности имеет такой же физический смысл, как и спектр плотности энергии. Пусть мощностной сигнал $f(t)$ действует на входе идеального полосового фильтра с передаточной функцией $H(\omega)$, показанной на рис. 2.10. Этот фильтр подавляет все частотные составляющие, за исключением узкой полосы $\Delta\omega\,(\Delta\omega\to 0)$ вблизи ω_0. Спектр плотности мощности на выходе равен $S_f(\omega_0)$ (рис. 2.14).

Мощность выходного сигнала

$$P_o = 2S_f(\omega_0)\,\Delta f. \tag{2.38}$$

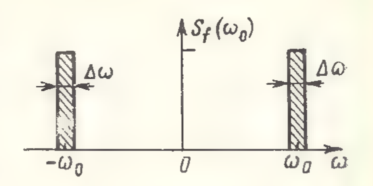


Рис. 2.14. Спектр плотности мощности узкополосного сигнала

Это не что иное, как мощность частотных составляющих сигнала $f(t)$, попадающих в полосу Δf вблизи ω_0. Мощности на положительных и отрицательных частотах одинаковы и равны $S_f(\omega_0)\Delta f$. Таким образом, $S_f(\omega)$ — мощность, приходящаяся на единицу полосы частот (*гц*) на частоте ω; она измеряется в *вт/гц*.

Пример 2.4

Найдем спектр плотности мощности производной.

Передаточная функция идеальной дифференцирующей цепи, как это следует из ф-лы (1.119), равна $i\omega$, поэтому

$$|H(\omega)|^2 = |i\,\omega|^2 = \omega^2$$

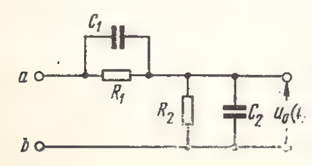
и спектр плотности мощности производной df/dt

$$S_{\dot f}(\omega) = \omega^2 S_f(\omega), \tag{2.39}$$

где $S_f(\omega)$ — спектр плотности мощности сигнала $f(t)$.

ЗАДАЧИ

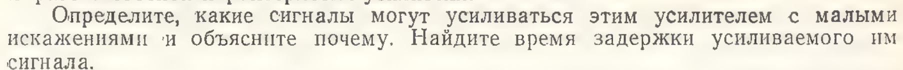
2.1. Покажите, что отклик идеального фильтра нижних частот, характеристики которого приведены на рис. 2.4, на сигнал $\frac{\pi}{W}\delta(t)$ совпадает с откликом на сигнал $Sa(Wt)$. Объясните этот результат.



2.2. Цепь из резисторов R_1 и R_2 с паразитными емкостями C_1 и C_2 используется для уменьшения напряжения, приложенного к клеммам *ab* (см. рисунок). Каково должно быть соотношение между сопротивлениями этих резисторов и емкостями, чтобы передача была неискаженной?

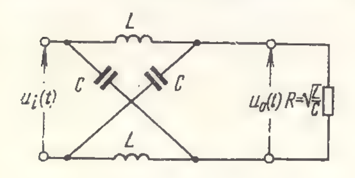
2.3. Найдите передаточную функцию RC-цепи, показанной на рисунке. Изобразите амплитудно-частотную и фазо-частотную характеристики цепи. Какие сигналы можно пропускать через эту цепь с достаточно малыми искажениями? Какова величина задержки выходного сигнала?

2.4. Усилитель имеет передаточную функцию $H(\omega)=K/(i\omega+\omega_0)$. Постройте графики амплитудно-частотной и фазо-частотной характеристик усилителя.



Определите, какие сигналы могут усиливаться этим усилителем с малыми искажениями и объясните почему. Найдите время задержки усиливаемого им сигнала.

2.5. Определите передаточную функцию скрещенной цепи, изображенной на рисунке. Найдите соотношение, связывающее выходное напряжение $u_0(t)$ со входным $u_i(t)$. Представьте амплитудно-частотную и фазо-частотную характеристики. Определите, можно ли использовать эту цепь для неискаженной передачи сигналов. Каким условиям должны удовлетворять сигналы, которые могут проходить через данную цепь с малыми искажениями? Найдите входное сопротивление цепи.



Определите, можно ли создать линию задержки, соединив последовательно большое число таких скрещенных цепей. При каком условии (налагаемом на сигнал) задержка сигнала не сопровождается большими искажениями? Какая задержка получается при последовательном включении n звеньев?

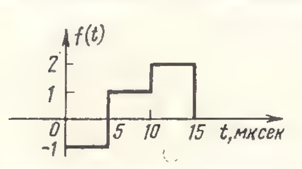
2.6. Найдите соотношение, связывающее выходное напряжение $u_0(t)$ со входным током $i(t)$, для цепи, изображенной на рисунке. Определите параметры элементов цепи, при которых выходное напряжение имеет ту же форму, что и входной ток. Возникает ли задержка при передаче сигнала?

2.7. Передаточная функция идеального полосового фильтра имеет вид:

$$H(\omega)=\kappa\left[G_W(\omega-\omega_0)+G_W(\omega+\omega_0)\right]e^{-i\omega t_0}.$$

Представьте амплитудно-частотную и фазо-частотную характеристики этого фильтра. Определите его импульсную реакцию. Изобразите график импульсной реакции и установите, является ли фильтр физически реализуемым.

2.8. Определите отклик полосового фильтра из задачи 2.7 на входной сигнал $\cos\omega_0 t u_1(t)$ (предположите, что полоса пропускания фильтра достаточно узка, чтобы можно было приближенно считать $1/[i(\omega+\omega_0)]\approx 1/(2i\omega_0)$ вблизи $\omega=\omega_0$ и $1/[i(\omega-\omega_0)]\approx -1/(2i\omega_0)$ вблизи $\omega=-\omega_0$.

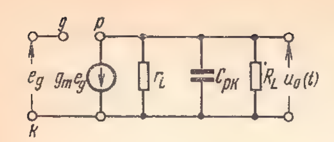


2.9. Определите и изобразите отклик идеального фильтра нижних частот (рис. 2.4*а*) на сигнал, показанный на рисунке (слева), если частота среза фильтра равна 10 *кгц*, время задержки $t_0=1$ *мксек*.

2.10. Найдите реакцию на единичный импульс и единичный скачок идеального фильтра верхних частот, передаточная функция которого

$$H(\omega)=\left[1-C_{2W}(\omega)\right]e^{-i\omega t_0}.$$

2.11. Определите передаточную функцию $H(\omega)$ и полосу пропускания лампового усилителя, эквивалентная схема которого показана на рисунке.

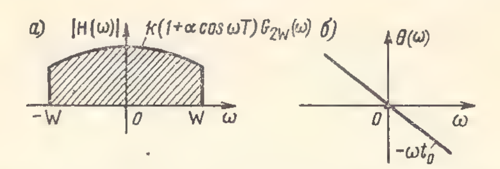


Постройте график выходного напряжения $u_0(t)$, если ко входу gk приложен единичный скачок напряжения.

Покажите, что произведение полосы пропускания (в *рад/сек*) на время нарастания равно постоянной величине, не зависящей от параметров лампы (время нарастания определяется здесь как интервал времени, в течение которого выходное напряжение при подаче на вход единичного скачка напряжения изменяется от 0,1 до 0,9 своего установившегося значения), и что эта постоянная равна 2,2.

Покажите, что произведение усиления постоянного тока $H(0)$ на полосу пропускания равно g_m/C_{pk}.

2.12. Допустим, что сигнал $f(t)$ ограничен по спектру, т. е. его спектр $F(\omega)=0$ для $(\omega)>W$. Этот сигнал можно усиливать без искажений усилителем, имеющим такую же передаточную функцию $H(\omega)$, как и идеальный фильтр нижних частот: $H(\omega)=kG_{2W}(\omega)e^{-i\omega t_0}$. Если амплитудно-частотная и фазо-частотная характеристики усилителя отклоняются от идеальных, то возникают искажения, известные как парные эхо. Пусть фазо-частотная характе-



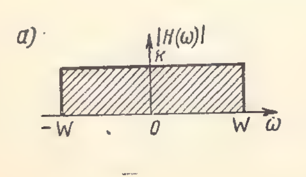
ристика усилителя — идеальная $\theta(\omega)=-\omega t_0$, а амплитудно-частотная имеет завал на верхних частотах, как это показано на рисунке:

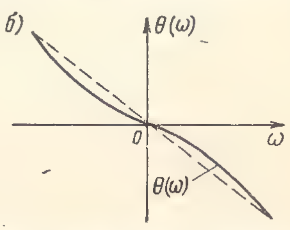
$$H(\omega)=\kappa(1+\alpha\cos\omega T)\,G_{2W}(\omega).$$

Найдите напряжение на выходе усилителя, если на его вход поступает импульсный сигнал $f(t)$, спектр которого ограничен полосой W, *рад/сек*.

Указание: используйте ф-лу (1.118).

2.13. Пусть усилитель из задачи 2.12 имеет идеальную амплитудно-частотную характеристику, т. е. $H(\omega)=kG_{2W}(\omega)$, а его фазо-частотная характеристика отличается от идеальной (см. рисунок): $\theta(\omega)=-\omega t_0+\alpha\sin\omega T$. Найдите выходное напряжение усилителя, если на его входе действует сигнал $f(t)$ из задачи 2.12.





Указание. Предположите, что α и T — весьма малые величины, и разложите $e^{i\alpha\sin\omega T}$ в ряд Тейлора, оставив два первых члена ряда:

$$e^{i\alpha\sin\omega T}\approx 1+i\alpha\sin\omega T=1+\frac{\alpha}{2}\left(e^{i\omega T}-e^{-i\omega T}\right).$$

2.14. Сигнал $f(t)=2e^{-t}u_1(t)$ проходит через идеальный фильтр нижних частот с частотой среза 1 *рад/сек*. Найдите спектр плотности энергии выходного сигнала. Определите энергии входного и выходного сигналов.

2.15. Для энергетического сигнала $f(t)$ покажите, что спектральная плотность энергии $\Psi_f(\omega)$ есть преобразование Фурье функции $\varphi(\tau)$, где

$$\varphi(\tau) = \int_{-\infty}^{\infty} f(t) f(t-\tau)\, dt = \int_{-\infty}^{\infty} f(t) f(t+\tau)\, dt.$$

2.16. Выведите теорему Парсеваля [ф-ла (2.13)] из теоремы временной свертки.

Указание. Если $f(t) \longleftrightarrow F(\omega)$, то $f(-t) \longleftrightarrow F(-\omega)$ и $f(t) * f(-t) \longleftrightarrow |F(\omega)|^2$.

2.17. Для действительного мощностного сигнала $f(t)$ покажите, что спектральная плотность мощности $S_f(\omega)$ является преобразованием Фурье функции $Rf(\tau)$, где

$$R_f(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t-\tau)\, dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t+\tau)\, dt.$$

Функция $R_f(\tau)$ известна как временная автокорреляционная функция сигнала $f(t)$.

2.18. Мощность сигнала определяется как мощность, рассеиваемая им на сопротивлении в 1 *ом*. Мощность действительного сигнала $f(t)$ определяется соотношением (2.18). Выведите подобную формулу для случая комплексного сигнала $f(t)$.

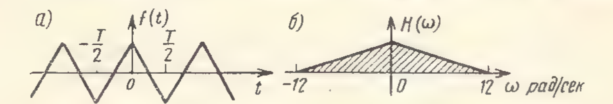
2.19. Определите мощность (как среднее значение квадрата сигнала) следующих сигналов и изобразите их спектры плотности мощности.

а) $A\cos(2000\pi t) + B\sin(200\pi t)$, г) $A\sin(200\pi t)\cos(2000\pi t)$,

б) $[A + \sin(200\pi t)]\cos(2000\pi t)$ д) $A\sin(300\pi t)\cos(2000\pi t)$,

в) $A\cos(200\pi t)\cos(2000\pi t)$, е) $A\sin^2(200\pi t)\cos(2000\pi t)$.

2.20. Показанный на рисунке периодический сигнал $f(t)$ проходит через цепь с передаточной функцией $H(\omega)$. Найдите спектр плотности мощности и сред-



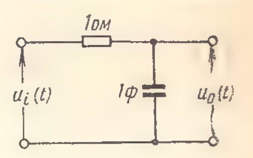
нюю мощность выходного сигнала для трех значений $T (T = 2\pi/3,\ \pi/3,\ \pi/6)$. Вычислите мощность входного сигнала $f(t)$.

2.21. Определите среднее значение квадрата сигнала $u_0(t)$ на выходе RC-цепи (см. рисунок), если спектр плотности мощности $S_i(\omega)$ входного сигнала:

а) $S_i(\omega) = K$;

б) $S_i(\omega) = G_2(\omega)$ (прямоугольная функция с частотой среза $\omega = 1$);

в) $S_i(\omega) = \pi[\delta(\omega + 1) - \delta(\omega - 1)]$.



Вычислите в каждом случае среднюю мощность входного сигнала.

ГЛАВА

3

Системы связи с амплитудной модуляцией

3.1. Временно́е и частотное уплотнение

Все системы связи основаны на передаче различных сигналов из одного пункта в другой. Эта задача встречается в радио и телевизионном вещании, связи по телефонным линиям на большие расстояния, спутниковой связи, системах телеуправления, телеметрии и т. д. Некоторые способы связи изучаются в этой главе.

Сигналы из одного пункта в другой передаются по каналу, который может быть парой проводов (например, телефонный) или просто открытым пространством, в которое излучаются сигналы, несущие полезную информацию (радио и телевизионное вещание, спутниковая связь и т. п.). Каждый из передаваемых сигналов обычно имеет малую ширину спектра по сравнению с полосой пропускания канала. Поэтому передавать лишь один сигнал по каналу невыгодно, так как емкость канала будет использоваться весьма незначительно. Вместе с тем непосредственная передача по каналу одновременно нескольких сигналов привела бы к наложению этих сигналов и выделение их на приемной стороне оказалось бы невозможным. Указанная трудность преодолевается посредством использования частотного или временно́го уплотнения.

Если сдвинуть частотные спектры различных сигналов так, чтобы они занимали неперекрывающиеся частотные полосы, можно передавать одновременно большое число сигналов по одному каналу. В гл. 1 было показано, что сдвиг частотного спектра выполняется посредством модуляции, т. е. умножения сигнала на синусоидальное колебание.

Спектр каждого сигнала сдвигается так, чтобы перекрытие спектров соседних сигналов отсутствовало. На приемной стороне сигналы разделяются частотными фильтрами. Однако, выделенный фильтром сигнал еще не есть исходное сообщение, так как его спектр сдвинут по частоте. Для того чтобы получить исходное сообщение, необходимо выполнить обратный перенос спектра сигнала.

В системах связи с излучением электромагнитных волн модуляция служит и для другой цели. В теории электромагнитных волн показывается, что эффективное излучение возможно лишь в том случае, если размер антенны составляет не менее одной десятой длины волны излучаемого сигнала. Речевой сигнал имеет наивысшую частоту около 10 *кгц,* что соответствует наименьшей длине волны 30 000 *м.* Таким образом, для излучения электромагнитных волн, соответствующих речевому сигналу, потребовалась бы антенна длиной в несколько километров, что конечно, непрактично. Модуляция сдвигает спектр сигнала на любую требуемую частоту, облегчая, таким образом, задачу излучения сигнала. Обычно несущая частота сигнала весьма высока[1]). Итак, модуляция не только решает задачу одновременной передачи нескольких сигналов, но и обеспечивает возможность их эффективного излучения.

Метод частотного сдвига сигналов — не единственный способ одновременной передачи нескольких сигналов по одному каналу. В гл. 1 показано, что сигнал с ограниченным спектром, в котором не содержатся составляющие с частотами, выше некоторой частоты f_m, *гц,* однозначно определяется своими мгновенными значениями (отсчетами), следующими через $1/(2f_m)$ *сек.* Эти отсчеты позволяют восстановить непрерывный сигнал. Следовательно, достаточно передавать лишь отсчеты сигнала. Свободные от передачи интервалы времени между отсчетами можно занять для передачи отсчетов других сигналов. На приемной стороне импульсы, принадлежащие каждому сигналу, выделяются синхронным коммутатором.

Таким образом, несколько сигналов могут передаваться по каналу одновременно при условии, что их разделение на приемной стороне возможно. Каждый сигнал определяется в частотной или во временно́й области. Поэтому и разделение сигналов на приемной стороне можно рассматривать как в частотной, так и во временно́й областях. При частотном разделении сигналы имеют неперекрывающиеся спектры, но смешаны во временно́й области. Разделение их осуществляется частотными фильтрами. Принцип одновременной передачи сигналов, занимающих неперекрывающиеся полосы частот, называется *частотным уплотнением.* При временно́м разделении отсчеты различных сигналов перемежаются, и их разделение выполняется синхронным коммутатором. Спектры всех импульсных сигналов занимают одну и ту же полосу частот. Такой способ одновременной передачи называется *временны́м уплотнением.* В этой и двух последующих главах рассматриваются различные системы связи, использующие указанные выше способы уплотнения.

---

[1]) В системах радиовещания выше 150 *кгц.* — *Прим. ред.*

3.2. Амплитудная модуляция с подавленной несущей (АМ-ПН)

ОБЩИЕ ПОЛОЖЕНИЯ

При этом способе передачи перенос спектра передаваемого сигнала осуществляется перемножением этого сигнала на синусоидальное колебание, частота которого равна той, на которую требуется перенести спектр. Из теоремы о модуляции [ф-ла (1.116a)] следует, что колебание $f(t)\cos\omega_c t$ имеет такой же спектр, как и

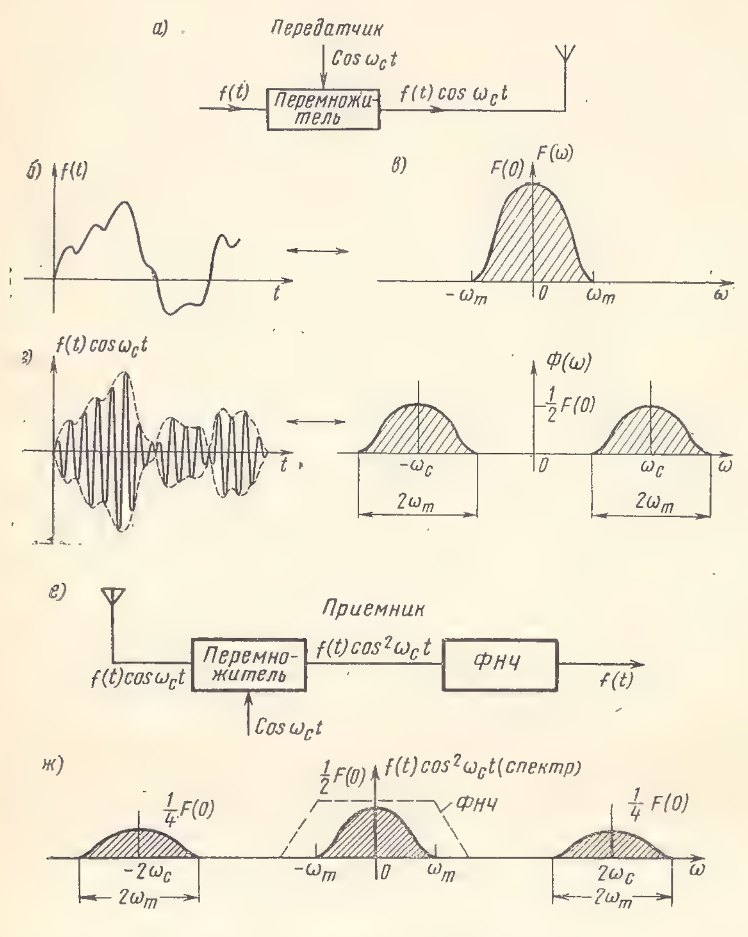


Рис. 3.1. Система связи с АМ-ПН:
а) передатчик; *б)* передаваемое сообщение $t(f)$; *в)* спектр сообщения $f(t)$; *г)* модулированный сигнал; *д)* спектр модулированного сигнала; *е)* приемник; *ж)* спектр сигнала на выходе перемножителя

сигнал $f(t)$, но сдвинутый на $\pm\omega_c$ *рад/сек* (рис. 3.1*д*), т. е. если $f(t)\longleftrightarrow F(\omega)$, то

$$f(t)\cos\omega_c t \longleftrightarrow \frac{1}{2}[F(\omega+\omega_c)+F(\omega-\omega_c)]. \qquad (3.1)$$

Колебание $\cos\omega_c t$ называется несущим. Умножение несущего колебания на сигнал $f(t)$ эквивалентно изменению его амплитуды пропорционально значениям $f(t)$. Таким образом, сигнал $f(t)$ есть модулирующий сигнал, а несущее колебание $\cos\omega_c t$ — модулируемый сигнал. Этот способ передачи называется амплитудной модуляцией с *подавленной несущей* (АМ-ПН) [1], так как в спектре модулированного сигнала $f(t)\cos\omega_c t$ не содержится составляющая несущей частоты (она оказывается подавленной).

При АМ-ПН частотный спектр передаваемого сообщения $f(t)$ перемещается на $\pm\omega_c$ *рад/сек*, как это следует из ф-лы (3.1). Чтобы получить исходное сообщение $f(t)$ из модулированного сигнала, необходимо сдвинуть спектр в его первоначальное положение. Процесс обратного переноса спектра в его первоначальное положение называется демодуляцией или детектированием.

Перенос спектра модулированного сигнала (рис. 3.1*д*) осуществляется умножением его на $\cos\omega_c t$ в приемнике. Поскольку умножение во временно́й области эквивалентно свертке спектров в частотной области, спектр получающегося колебания $f(t)\cos^2\omega_c t$ находится как свертка спектра принимаемого сигнала (рис. 3.1*д*) со спектром колебания $\cos\omega_c t$ (два единичных импульса на частотах $\pm\omega_c$). Результатом свертки является спектр, показанный на рис. 3.1*ж* (см. пример 1.12, рис. 1.32). Этот результат также получается непосредственно из тождества

$$f(t)\cos^2\omega_c t = \frac{1}{2}f(t)[1+\cos 2\omega_c t] = \frac{1}{2}[f(t)+f(t)\cos 2\omega_c t]. \qquad (3.2)$$

Следовательно, если $f(t)\longleftrightarrow F(\omega)$, то

то

$$f(t)\cos^2\omega_c t \longleftrightarrow \frac{1}{2}F(\omega)+\frac{1}{4}[F(\omega+2\omega_c)+F(\omega-2\omega_c)]. \qquad (3.3)$$

Из рассмотрения спектра на рис. 3.1*ж* следует, что первоначальный сигнал $f(t)$ можно выделить, используя фильтр нижних частот, который пропускает спектр $F(\omega)$ и задерживает остальные составляющие вблизи частот $\pm 2\omega_c$.

Возможная форма характеристики фильтра нижних частот показана пунктирной линией на рис. 3.1*ж*. На рис. 3.1*е* показана

---

[1]) Ниже будет рассматриваться амплитудная модуляция с дополнительным несущим колебанием, при которой модулированный сигнал равен $f(t)\cos\omega_c t + A\cos\omega_c t$. Передача дополнительного несущего колебания имеет ряд преимуществ. Такие системы называются системами амплитудной модуляции (АМ).

система, выделяющая сообщение $f(t)$ из принимаемого модулированного сигнала $f(t)\cos\omega_c t$. Интересно отметить, что процесс обработки сигнала на приемной стороне такой же, как на передающей. Рассмотренный способ выделения сообщения из модулированного сигнала называется синхронным или когерентным детектированием.

Из изложенного следует, что при такой системе передачи необходимо генерировать местную несущую в приемнике. Частота и фаза местной несущей крайне критичны. Например, рассмотрим местную несущую с небольшой частотной ошибкой $\Delta\omega$. Принимаемый сигнал равен $f(t)\cos\omega_c t$, а местная несущая равна $\cos(\omega_c+\Delta\omega)t$. Перемножив их, получим

$$f(t)\cos\omega_c t\cos(\omega_c+\Delta\omega)\,t=\frac{1}{2}f(t)[\cos\Delta\omega\, t+\cos(2\omega_c+\Delta\omega)t\,]. \quad (3.4)$$

Слагаемое $f(t)\cos(2\omega_c+\Delta\omega)t$ представляет собой спектр $f(t)$, перенесенный на частоты $\pm(2\omega_c+\Delta\omega)$; он может быть отфильтрован фильтром нижних частот. На выходе фильтра напряжение соответствует слагаемому $\frac{1}{2}f(t)\cos\Delta\omega t$. Таким образом, вместо сообщения $f(t)$ получаем $f(t)\cos\Delta\omega t$. В общем случае отклонение частоты мало́ ($\Delta\omega\to 0$), и $f(t)\cos\Delta\omega t$ есть сообщение $f(t)$, умноженное на медленно изменяющуюся функцию. Очевидно, искажение такого рода является недопустимым. Поэтому крайне важно иметь одинаковые несущие частоты в передатчике и приемнике. Фаза местной несущей также критична. Более подробно этот вопрос рассматривается в § 3.5.

Для точной регулировки частоты и фазы местной несущей в приемнике требуются очень сложные и дорогостоящие устройства. В большинстве систем вместе с модулированным сигналом передается незначительная часть несущей (пилот-сигнал). В приемнике пилот-сигнал отфильтровывается и усиливается. Полученное колебание затем используется для автоподстройки фазы местного генератора, который генерирует несущее колебание большого уровня и такой же частоты, как в передатчике.

МЕТОДЫ ПРЕОБРАЗОВАНИЯ ЧАСТОТЫ

Из теоремы о модуляции следует, что спектр любого сигнала можно сдвинуть в частотной области на $\pm\omega_c$ *рад/сек*, умножив его на гармоническое колебание частоты ω_c. Это, однако, не единственный способ. Легко показать, что спектр сигнала можно сдвинуть на частоты $\pm\omega_c$, умножив его на произвольный периодический сигнал частоты ω_c. Действительно, любое периодическое колебание частоты ω_c содержит гармонические составляющие частот 0, ω_c, $2\omega_c$, $3\omega_c$ ... Следовательно, при умножении сигнала $f(t)$ на произвольное периодическое колебание частоты ω_c спектр $f(t)$ перемещается на частоты 0, $\pm\omega_c$, $\pm 2\omega_c$ ... и т. д. Интерес, однако,

представляет лишь та часть спектра, которая расположена вблизи $\pm\omega_c$. Ее можно выделить полосовым фильтром, который пропускает частотные составляющие вблизи $\pm\omega_c$, а все остальные подавляет.

Рассмотрим, например, сигнал $f(t)$ и его спектр $F(\omega)$ (рис. 3.2*а*). При умножении этого сигнала на гармоническое колебание $\cos\omega_c t$ (рис. 3.2*б*) спектр $F(\omega)$ сдвигается на частоты $\pm\omega_c$ (рис. 3.2*в*). Пусть теперь сигнал $f(t)$ умножается на прямоугольное колебание $p(t)$ частоты ω_c (рис. 3.2*г*). Спектр $P(\omega)$ этого колебания представляет собой последовательность единичных импульсов на частотах $\omega=0$, $\pm\omega_c$, $\pm 3\omega_c$, $\pm 5\omega_c$ ... (см. также рис. 1.23.) Очевидно, спектр $f(t)p(t)$ равен $(1/2\pi)F(\omega) * P(\omega)$. Результат этой свертки показан на рис. 3.2*д*.

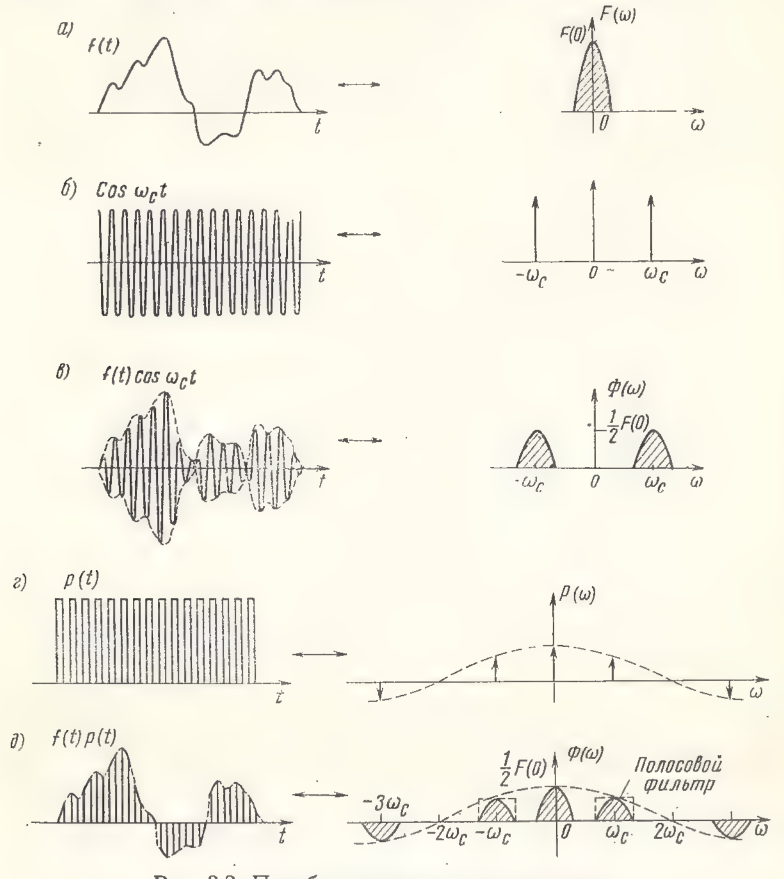


Рис. 3.2. Преобразование частоты сигнала:
а) сигнал $f(t)$ и его спектр; *б)* гармоническое колебание и его спектр; *в)* сигнал $f(t)\cos\omega_c t$ и его спектр; *г)* периодический сигнал прямоугольной формы $p(t)$ и его спектр; *д)* сигнал $f(t)p(t)$ и его спектр

Итак, умножение $f(t)$ на $p(t)$ сдвигает спектр сигнала $f(t)$ на частоты $\omega=0$, $\pm\omega_c$, $\pm 3\omega_c$, $\pm 5\omega_c$... Этот результат справедлив для любой периодической функции частоты ω_c независимо от ее формы. В частном случае прямоугольного колебания, у которого длительности импульсов и пауз одинаковы, четные гармоники $\pm 2\omega_c$, $\pm 4\omega_c$... равны нулю. В общем же случае это не обязательно. Поэтому можно сделать вывод, что умножение сигнала $f(t)$ на любой периодический сигнал частоты ω_c, независимо от его формы, сдвигает спектр на частоты $\omega=0$, $\pm\omega_c$, $\pm 2\omega_c$, $\pm 3\omega_c$ ...

Этот результат легко получить аналитически. Пусть $\varphi(t)$ — периодический сигнал частоты f_c, *гц* ($\omega_c=2\pi f_c$). В общем случае преобразование Фурье сигнала [ф-ла (1.102)] записывается как

$$\varphi(t) \longleftrightarrow 2\pi \sum_{n=-\infty}^{\infty} \Phi_n \delta(\omega - n\omega_c), \tag{3.5}$$

где Φ_n — коэффициенты Фурье для n-й гармоники.

Из теоремы свертки следует, что

$$f(t)\varphi(t) \longleftrightarrow \frac{1}{2\pi} F(\omega) * 2\pi \sum_{n=-\infty}^{\infty} \Phi_n \delta(\omega - n\omega_c) \tag{3.6}$$

$$\longleftrightarrow \sum_{n=-\infty}^{\infty} \Phi_n F(\omega) * \delta(\omega - n\omega_c)$$

$$\longleftrightarrow \sum_{n=-\infty}^{\infty} \Phi_n F(\omega - n\omega_c). \tag{3.7}$$

Из ф-лы (3.7) следует, что спектр произведения $f(t)\varphi(t)$ состоит из спектра $F(\omega)$ и аналогичных спектров, сдвинутых на частоты $\pm\omega_c$, $\pm 2\omega_c$ ... Заметим, что значения спектральной плотности на всех этих частотах находятся умножением $F(\omega)$ на постоянные числа Φ_0, Φ_1, Φ_2 ...

В случае прямоугольного колебания $\varphi(t)$ коэффициенты Φ_n находятся по ф-ле (1.66а) подстановкой в нее $T=2\tau$ и $A=1$:

$$\Phi_n = \frac{1}{2} Sa\left(\frac{n\pi}{2}\right).$$

Заметим, что

$$Sa\left(\frac{n\pi}{2}\right) = \frac{\sin(n\pi/2)}{n\pi/2} = \begin{cases} (-1)^{(n-1)/2}\left(\frac{2}{n\pi}\right) & \text{при } n \text{ нечетном,} \\ 1 & \text{при } n=0, \\ 0 & \text{при } n \text{ четном.} \end{cases}$$

Следовательно, из ф-лы (3.5) получаем:

$$p(t) \longleftrightarrow \pi\delta(\omega) + 2 \sum_{\substack{n=-\infty \\ (n=3,5\ldots)}}^{\infty} \frac{(-1)^{(n-1)/2}}{n} \delta(\omega - n\omega_c), \tag{3.8}$$

$$f(t)\,p(t) \longleftrightarrow \frac{1}{2} F(\omega) + \frac{1}{\pi} \sum_{\substack{n=-\infty \\ (n=3,5\ldots)}}^{\infty} \frac{(-1)^{(n-1)/2}}{n} F(\omega - n\omega_c). \tag{3.9}$$

Найденный по ф-ле (3.9) спектр показан на рис. 3.2*д*.

Однако в случае амплитудной модуляции интерес представляет лишь часть спектра вблизи $\pm\omega_c$. Ее можно выделить полосовым фильтром. Обычный резонансный контур *RLC*, настроенный на частоту ω_c, будет пропускать полосы частот вблизи $\pm\omega_c$ и отфильтровывать остальные составляющие. Следовательно, если пропустить сигнал $f(t)p(t)$ (рис. 3.2*д*) через такой полосовой фильтр, то на выходе последнего образуется колебание $f(t)\cos\omega_c t$, как показано на рис. 3.3.

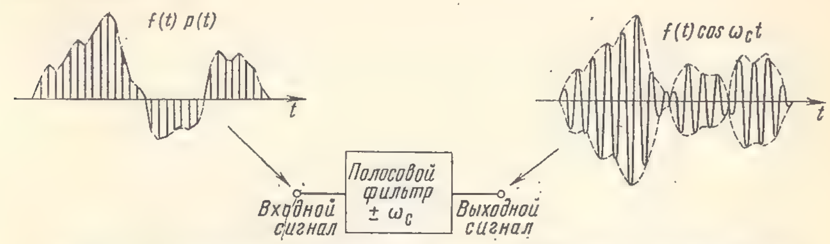


Рис. 3.3. К пояснению фильтрации сигнала $f(t)\cos\omega_c t$ из колебания $f(t)p(t)$

Процесс переноса спектра называется также *преобразованием частоты*, а цепи, которые его выполняют, — *преобразователями частоты*. Как модулятор, так и детектор в рассматриваемом случае выполняют операцию переноса спектра и, следовательно, являются преобразователями частоты.

При обсуждении методов переноса спектра часто будут упоминаться различные фильтры: нижних частот, верхних частот, полосовой. Можно получить фильтр с амплитудно-частотной (или фазо-частотной) характеристикой, весьма близкой к идеальной, применив большое число звеньев. Но в большинстве случаев нежелательные частотные составляющие, которые должны быть отфильтрованы, отстоят далеко от полезных составляющих, и поэтому можно использовать простые фильтры.

АМПЛИТУДНЫЕ МОДУЛЯТОРЫ (ПРЕОБРАЗОВАТЕЛИ ЧАСТОТЫ)

Рассмотрим некоторые простые устройства, при помощи которых осуществляется модуляция. В процессе модуляции происходит перенос спектра. Следовательно, выходной сигнал модулятора

должен содержать частоты, отличающиеся от частот входного сигнала. Поэтому получить модуляцию в линейных цепях с постоянными параметрами невозможно, так как отклик этих цепей не может содержать частоты, отличные от частот входного сигнала. Для модуляции можно использовать линейные цепи с переменными параметрами (такие, как прерыватель), а также цепи с нелинейными элементами. Именно благодаря нелинейности возникает модуляция, но очень часто выполняющая модуляцию цепь может рассматриваться как линейная цепь с переменными параметрами.

Схема модулятора с использованием прерывателя показана на рис. 3.4*а*. Переключатель *s* поочередно подключается к зажимам *a* и *b* с частотой ω_c. Одну половину периода зажим *c* подключен

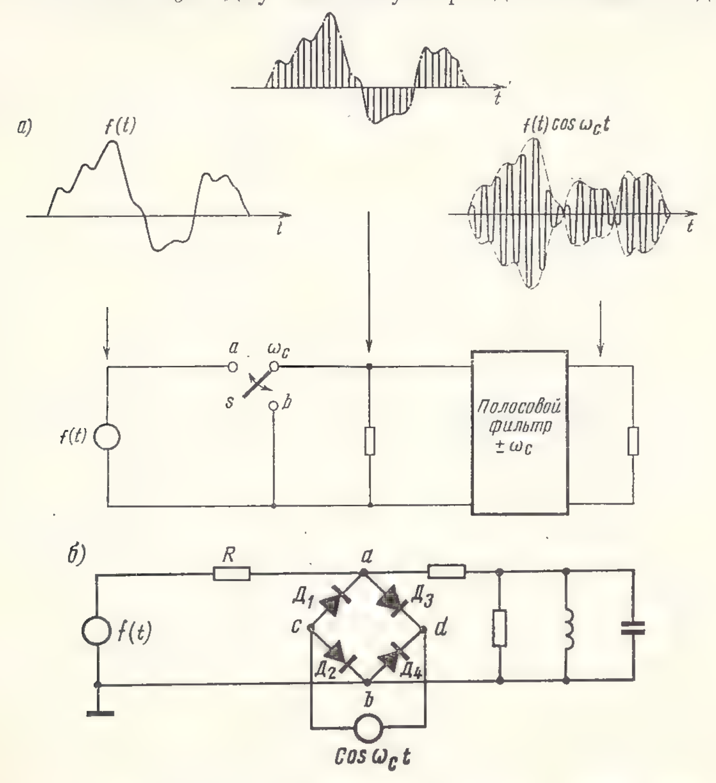


Рис. 3.4. Модулятор-прерыватель для получения сигналов АМ-ПН: *а*) схема, поясняющая его принцип действия; *б*) балансный (кольцевой) модулятор с диодным ключом

к источнику сообщения $f(t)$, другую — заземлен. Поэтому сигнал в точке *c* является прерывистым с частотой ω_c. Прерывание можно рассматривать как умножение сообщения $f(t)$ на прямоугольное колебание $p(t)$. Как указывалось выше, такое прерыви-

стое колебание содержит частотные составляющие на $\omega=0$, $\pm\omega_c$, $\pm 3\omega_c$ ... и полезный модулированный сигнал $f(t)\cos\omega_c t$ из этого колебания можно выделить, пропустив его через полосовой фильтр, настроенный на частоту ω_c (рис. 3.3).

Практическое осуществление указанного способа иллюстрируется рис. 3.4*б*. Диоды в этой схеме действуют подобно переключателю. Когда колебание $\cos\omega_c t$ имеет такую полярность, что в точке *c* напряжение положительно относительно точки *d*, все диоды открыты. Предполагается, что сигнал $\cos\omega_c t$ значительно превосходит колебание $f(t)$. При этом условии падение напряжения на диоде $Д_1$ равно падению напряжения на диоде $Д_2$ и, следовательно, точка *a* имеет такой же потенциал, как точка *b*. Таким образом, выходной зажим *a* оказывается заземленным. Когда же полярность сигнала $\cos\omega_c t$ такова, что напряжение в точке *d* положительно относительно точки *c*, диоды запираются, а зажимы *a* и *b* размыкаются. При этом зажим *a* оказывается соединенным с источником сообщения $f(t)$ через резистор *R*.

Таким образом, этот диодный ключ периодически с частотой ω_c заземляет зажим *a*. К выходным зажимам подключен параллельный резонансный контур, настроенный на частоту ω_c, который выполняет роль полосового фильтра. Выходное напряжение контура и есть полезный модулированный сигнал, пропорциональный сигналу $f(t)\cos\omega_c t$. Заметим, что рассмотренный модулятор является линейной цепью, так как умножение сигнала $f(t)$ на постоянную величину изменяет его пропорционально этой постоянной. Вместе с тем, модулятор есть цепь с переменными параметрами, так как его параметры периодически изменяются. Показанный на рис. 3.4*б* модулятор называется кольцевым.

В общем случае линейный модулятор можно рассматривать как цепь, у которой коэффициент передачи изменяется во времени пропорционально подводимому сигналу $f(t)$: $G=Kf(t)$. Несущее колебание $\cos\omega_c t$ поступает ко входным зажимам цепи (рис. 3.5*а*). Очевидно, на выходе получается модулированный сигнал $Kf(t)\cos\omega_c t$. С другой стороны, для изменения коэффициента передачи можно использовать несущее колебание (рис. 3.5*б*), а сигнал $f(t)$ подавать ко входным зажимам. Рассмотренный модулятор с диодным переключателем относится к последнему типу. Он действует как система, коэффициент передачи которой периодически изменяется от нуля до единицы с частотой ω_c. Коэффициент передачи изменяется во времени по прямоугольному закону, а

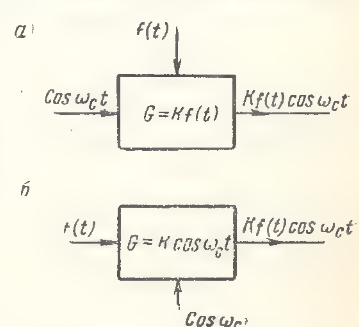


Рис. 3.5. Амплитудный модулятор как цепь с переменным параметром:
а) усиление цепи изменяется пропорционально сообщению $f(t)$; *б*) усиление цепи изменяется пропорционально несущей $\cos\omega_c t$

не по синусоидальному, что приводит к появлению нежелательных высших гармоник частоты ω_c, которые отфильтровываются.

В практических активных элементах, таких, как электронная лампа или транзистор, коэффициент усиления зависит от напряжения смещения и тока. Таким образом, коэффициент усиления этих элементов можно изменять во времени, изменяя под действием сигнала величину напряжения смещения. Подробное описание таких модуляторов (и детекторов) на электронных лампах и транзисторах можно найти в литературе по электронным цепям [1]).

Как отмечалось выше, модуляцию можно осуществить, используя нелинейные элементы. Характеристика типичного нелинейного элемента показана на рис. 3.6*а*. Примером такого элемента являет-

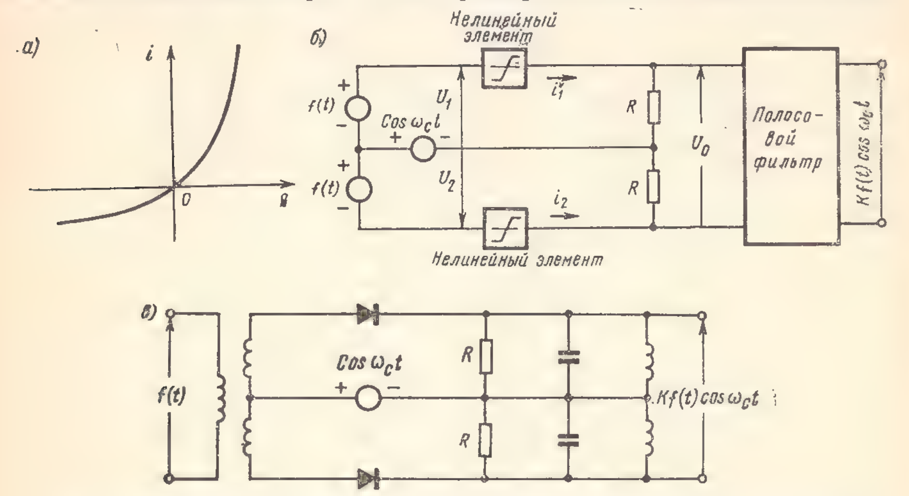


Рис. 3.6. Модулятор АМ-ПН с использованием нелинейных элементов: *а)* вольтамперная характеристика нелинейного элемента; *б)* схема, поясняющая принцип действия модулятора; *в)* балансный модулятор

ся полупроводниковый диод. Нелинейную характеристику подобного вида можно аппроксимировать степенным рядом [2]):

$$i = au + bu^2. \tag{3.10}$$

Подобная зависимость тока от напряжения наблюдается в транзисторах и электронных лампах при большом уровне сигнала. Одна из возможных схем модулятора с нелинейными элементами показана на рис. 3.6*в*. При анализе этой схемы будем рассматривать

[1]) См., например, А. М. Заездный и др. Теория нелинейных электрических цепей. М., «Связь», 1968. — *Прим. ред.*

[2]) Более точной является аппроксимация характеристики полупроводникового диода $i = a\,(e^{bu} - 1)$, где a и b — параметры аппроксимации, однако нахождение спектра тока оказывается более простым при аппроксимации степенным полиномом. — *Прим. ред.*

нелинейный элемент и последовательно включенный резистор R как составной нелинейный элемент, для которого входное напряжение u и ток i связаны соотношением (3.10).

Напряжения u_1 и u_2 (рис. 3.6*б*):

$$u_1 = \cos\omega_c t + f(t), \quad u_2 = \cos\omega_c t - f(t).$$

Очевидно, токи i_1 и i_2:

$$i_1 = au_1 + bu_1^2 = a[\cos\omega_c t + f(t)] + b[\cos\omega_c t + f(t)]^2, \qquad (3.10а)$$

$$i_2 = a[\cos\omega_c t - f(t)] + b[\cos\omega_c t - f(t)]^2. \qquad (3.10б)$$

Выходное напряжение u_o определяется как $u_o = i_1R - i_2R$. Подстановка в это выражение токов дает

$$u_o(t) = 2R\,[2bf(t)\cos\omega_c t + af(t)].$$

Слагаемое $2Raf(t)$ может быть отфильтровано с помощью полосового фильтра, настроенного на ω_c. В качестве нелинейных элементов в этой схеме удобно использовать полупроводниковые диоды. Практическая схема такого модулятора показана на рис. 3.6*в*. Все рассмотренные модуляторы вырабатывают сигнал амплитудной модуляции с подавленной несущей и называются *балансными*.

ДЕТЕКТИРОВАНИЕ МОДУЛИРОВАННЫХ СИГНАЛОВ С ПОДАВЛЕННОЙ НЕСУЩЕЙ

Для восстановления на приемной стороне передаваемого сообщения $f(t)$ необходимо продетектировать принимаемый сигнал $f(t)\cos\omega_c t$. Как указывалось выше, процесс детектирования эквивалентен переносу спектра и может выполняться перемножением модулированного сигнала $f(t)\cos\omega_c t$ на колебание $\cos\omega_c t$ (синхронное детектирование). Поэтому детектирование можно осуществить, используя ту же схему, которая применялась для модуляции. Однако между схемами модуляции и детектирования имеется различие. Спектр на выходе модулятора находится вблизи частот $\pm\omega_c$, поэтому на выходе модулятора включается полосовой фильтр, настроенный на частоту ω_c. В случае детектора выходной спектр есть $F(\omega)$ и сосредоточен он вблизи $\omega=0$. Следовательно, на выходе необходимо включить фильтр нижних частот, чтобы отфильтровать нежелательные высокочастотные составляющие, спектры которых находятся вблизи частот $\pm\omega_c$, $\pm 2\omega_c$, $\pm 3\omega_c$... Детекторы с диодным ключом и нелинейными элементами показаны на рис. 3.7*а* и *б* соответственно. В каждом из детекторов роль фильтра нижних частот выполняет цепь RC.

Детектирование может осуществляться перемножением модулированного сигнала $f(t)\cos\omega_c t$ на любой периодический сигнал частоты ω_c. Если $\varphi(t)$ — периодический сигнал частоты ω_c, то его преобразование Фурье $\Phi(\omega)$ можно записать в виде ф-лы (3.5).

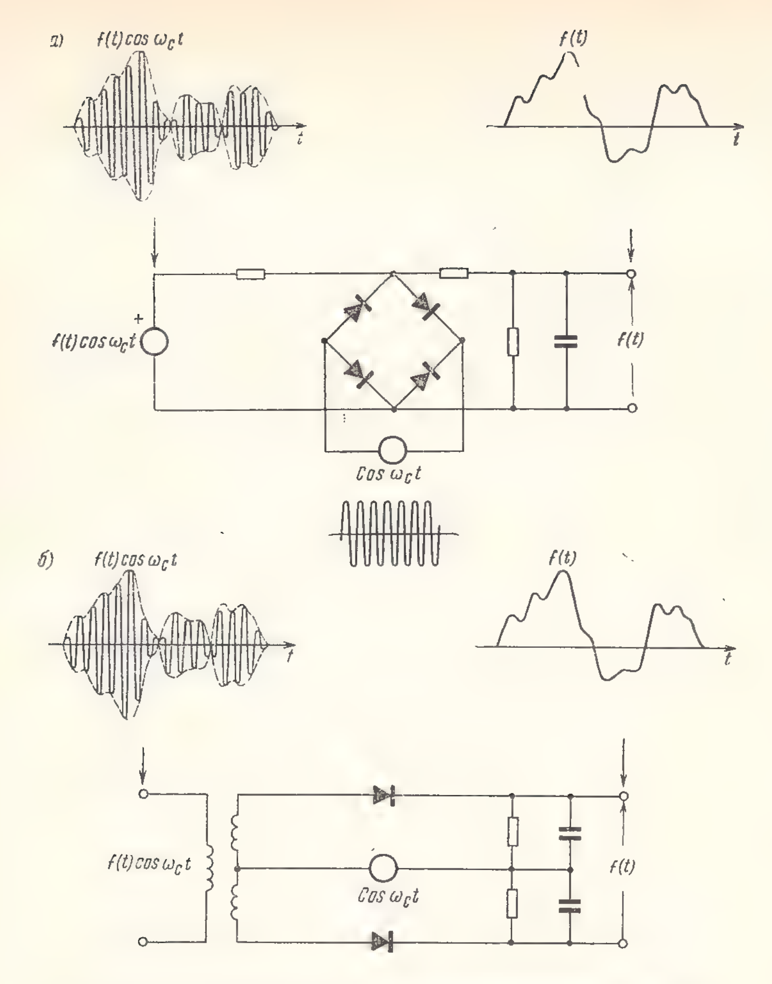


Рис. 3.7. К пояснению детектирования сигналов АМ-ПН:
а) балансный (кольцевой) детектор; *б)* балансный детектор
с нелинейными элементами

Если модулированный сигнал $f(t)\cos\omega_c(t)$ умножается на $\varphi(t)$, то спектр произведения

$$f(t)\cos\omega_c t\,\varphi(t) \longleftrightarrow \pi\,[F(\omega-\omega_c)+F(\omega+\omega_c)] * \sum_{n=-\infty}^{\infty}\Phi_n\delta(\omega - n\,\omega_c) \longleftrightarrow \pi\sum_{n=-\infty}^{\infty}\Phi_n\{F[\omega-(n+1)\,\omega_c]+F[\omega-(n-1)\,\omega_c]\}. \quad (3.11)$$

Этот спектр содержит спектр $F(\omega)$, который может быть выделен фильтром нижних частот.

Принцип преобразования частоты сигнала находит применение также в усилителях постоянного тока и усилителях низкочастотных сигналов. Трудно создать усилители на очень низких частотах из-за чрезвычайно больших размеров конденсаторов межкаскадной связи. Поскольку для токов низких частот конденсатор должен представлять короткое замыкание, размеры переходных конденсаторов в многокаскадных усилителях оказываются очень большими. Поэтому при усилении сигналов постоянного тока и сигналов сверхнизких частот используют непосредственное включение каскадов усилителя. При этом, однако, возникает серьезная проблема дрейфа рабочей точки усилителя. Обусловленные дрейфом паразитные изменения тока накладываются на выходной сигнал, причем эти изменения тока нельзя отличить от тех, которые создаются полезным входным сигналом.

Задача устранения вредного влияния дрейфа успешно решается с помощью усилителя-прерывателя, который, по существу, сдвигает спектр входного сигнала из области низких частот на соответствующую высокую частоту, где он может быть легко усилен. Затем усиленный высокочастотный сигнал детектируется для получения низкочастотного сигнала в усиленном виде. В качестве усилителя-прерывателя можно использовать любую схему из рассмотренных выше. Однако обычно для модуляции и детектирования применяют механический прерыватель, который периодически заземляет входные и выходные клеммы усилителя. Поскольку для модуляции и детектирования требуется одна и та же несущая частота, необходимо использовать один прерыватель, как показано на рис. 3.8*а*.

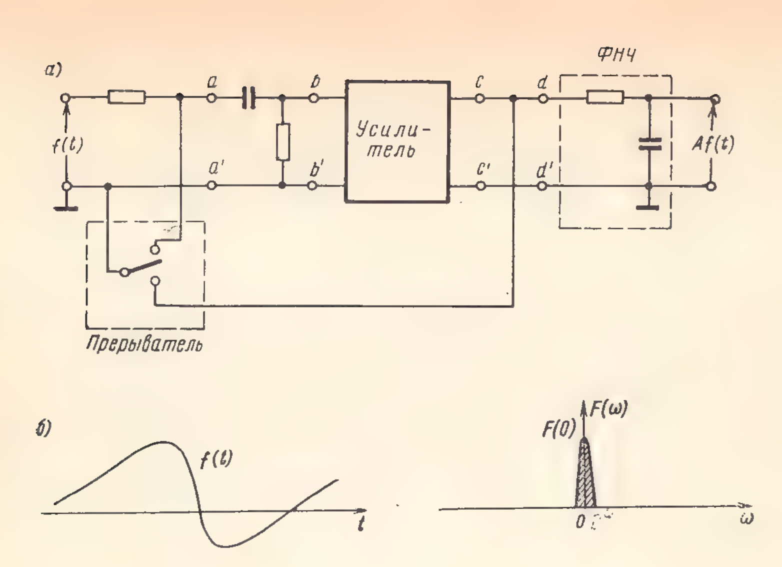
Рассмотрим низкочастотный сигнал $f(t)$ и его спектр $F(\omega)$ (рис. 3.8*б*). Сигнал $f(t)$ поступает на входные зажимы усилителя, как показано на рис. 3.8*а*. Каждые полпериода прерыватель заземляет $f(t)$. Возникающий на зажимах aa' прерывистый сигнал показан на рис. 3.8*в*. Этот сигнал есть входной сигнал $f(t)$, умноженный на прямоугольное колебание $p(t)$. Спектр прерывистого сигнала на зажимах aa' обозначен $F_{aa'}(\omega)$ и показан на рис. 3.8*в* (см. также рис. 3.2*д*).

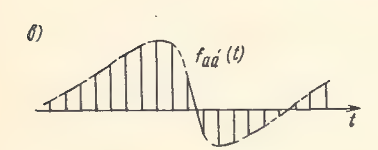
Входной конденсатор не пропускает часть спектра вблизи $\omega=0$. Поэтому спектр в точках bb' $F_{bb'}(\omega)$ аналогичен спектру $F_{aa'}(\omega)$, у которого отсутствуют составляющие вблизи частоты $\omega=0$. Спектр $F_{bb'}(\omega)$ показан на рис. 3.8*г*. Сигнал в точках bb' обозначается как $f_{bb'}(t)$ и равен обратному преобразованию Фурье спектра $F_{bb'}(\omega)$. Так как $F_{bb'}(\omega)=F_{aa'}(\omega)-\frac{1}{2}F(\omega)$, то $f_{bb'}(t)=$ $=f_{aa'}(t)-\frac{1}{2}f(t)$.

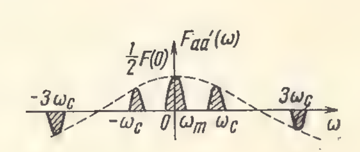
Таким образом, сигнал $f_{bb'}(t)$ можно получить, если вычесть половину исходного сигнала $f(t)$ из прерывистого сигнала $f_{aa'}(t)$. Сигнал $f_{bb'}(t)$ показан на рис. 3.8*г*[1]). Полученный результат очевиден, так как входной конденсатор не пропускает среднее значение сигнала $f_{aa'}(t)$. При этом сигнал преобразуется в биполярную форму $f_{bb'}(t)$. Сигнал $f_{bb'}(t)$ является входным сигналом усилителя. Заметим, что $f_{bb'}(t)$ уже не содержит составляющие очень низких частот и, следовательно, легко может быть усилен. Выходной сигнал усилителя показан на рис. 3.8*д*. Теперь этот сигнал детектируется тем же самым прерывателем. Прерыватель заземляет выход усилителя каждые полпериода. Следует отметить, что полупериоды входного и выходного напряжений дополняют друг друга.

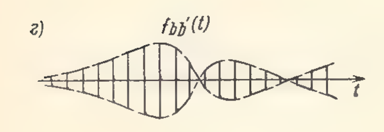
---

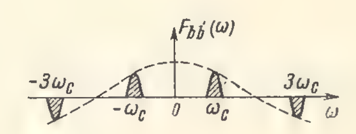
[1]) В действительности импульсы, указанные на рис. 3.8*г*, имеют несколько скошенные вершины, так как конденсатор не полностью подавляет низкочастотные составляющие. Однако, если частота ω_m очень мала, то скалывание вершин импульсов незначительно.

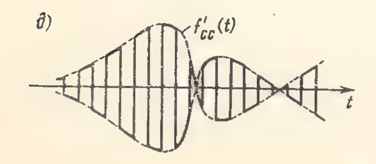


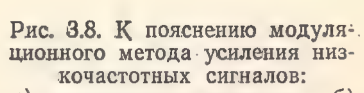














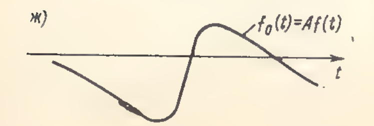


Рис. 3.8. К пояснению модуляционного метода усиления низкочастотных сигналов:

а) усилитель-прерыватель; *б)* низкочастотный сигнал и его спектр; *в)* сигнал на выходе прерывателя и его спектр; *г)* сигнал на выходе усилителя и его спектр; *д)* сигнал на выходе усилителя; *е)* сигнал на выходе демодулятора-прерывателя; *ж)* выходной усиленный сигнал

Продетектированный сигнал, который появляется на зажимах dd', обозначается $f_{dd'}(t)$ и показан на рис. 3.8*е*. Этот сигнал есть усиленный прерывистый сигнал $f_{aa'}(t)$. Следовательно, можно выделить сигнал $f(t)$, пропустив прерывистый сигнал $f_{dd'}(t)$ через фильтр нижних частот (цепь RC), показанный на рис. 3.8*а*. Окончательный выходной сигнал оказывается перевернутым. Однако большинство усилителей также вносит дополнительный фазовый сдвиг в 180°. В этих случаях выходной сигнал является усиленным колебанием $f(t)$ без изменения знака.

3.3. Амплитудная модуляция с большим уровнем несущей (АМ)

Как уже отмечалось, приемники сигналов АМ с подавленной несущей нуждаются в сложных устройствах синхронизации, необходимых для когерентного детектирования сигнала. Но такой способ передачи эффективен с точки зрения необходимой мощности передатчика. В одноканальных системах связи, где имеется один передатчик для каждого приемника, сложность последнего может быть оправдана в том случае, если она обеспечивает значительное снижение стоимости передатчика. С другой стороны, в системе радиовещания, где один передатчик работает на большое число приемников, более экономично иметь дорогостоящий передатчик и простейшие дешевые приемники. В таких системах связи вместе с сигналом $f(t)\cos\omega_c t$ передается несущее колебание большого уровня, что исключает необходимость генерации несущей в приемнике. Следовательно, в рассматриваемом случае передаваемый сигнал

$$\varphi_{AM}(t) = f(t)\cos\omega_c t + A\cos\omega_c t = [A + f(t)]\cos\omega_c t. \tag{3.12}$$

Очевидно, спектр этого сигнала отличается от спектра колебания $f(t)\cos\omega_c t$ лишь двумя единичными импульсами на частотах $\pm\omega_c$:

$$\varphi_{AM}(t) \longleftrightarrow \frac{1}{2}[F(\omega+\omega_c) + F(\omega-\omega_c)] + \pi A[\delta(\omega+\omega_c) + \delta(\omega-\omega_c)]. \tag{3.13}$$

Модулированный сигнал $\varphi_{AM}(t)$ показан на рис. 3.9*б*. Огибающая модулированного сигнала есть колебание $f(t)$, смещенное на постоянную величину A, поэтому выделение сигнала $f(t)$ сводится к детектированию огибающей. Заметим, что величина A должна быть достаточно большой, чтобы огибающая повторяла колебание $f(t)$. Если уровень A недостаточно велик (рис. 3.9*в*), то огибающая модулированного сигнала отличается от сигнала $f(t)$ [1], и последний нельзя извлечь из $\varphi_{AM}(t)$ посредством простого детектирования огибающей. В этом случае применяют синхронное детектирование (умножение модулированного сигнала на $\cos\omega_c t$). Следова-

[1]) Это явление называется перемодуляцией. — *Прим. ред.*

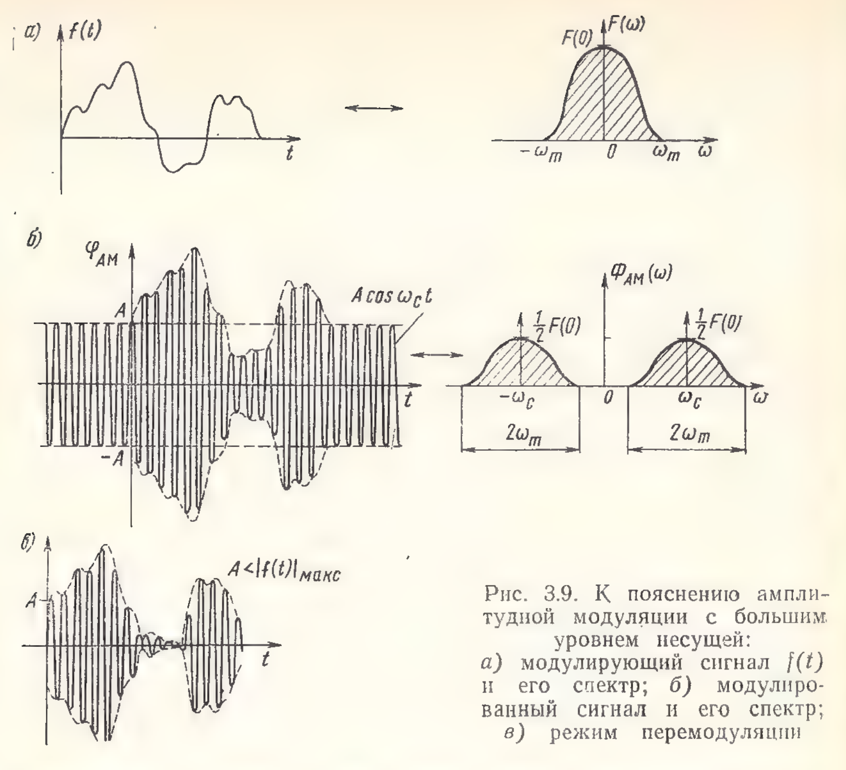


Рис. 3.9. К пояснению амплитудной модуляции с большим уровнем несущей:
а) модулирующий сигнал $f(t)$ и его спектр; *б)* модулированный сигнал и его спектр; *в)* режим перемодуляции

тельно, величина A должна быть такой, чтобы значение $[A+f(t)]$ оставалось всегда положительным. Это возможно, если

$$A > |f(t)|_{\text{макс}}. \tag{3.14}$$

Способ детектирования огибающей рассматривается ниже.

Модулированные сигналы, которые содержат несущее колебание большого уровня, удовлетворяющее условию (3.14), называются амплитудномодулированными (АМ). Ниже будет показано, что сигналы АМ значительно проще получить и детектировать, чем сигналы АМ-ПН.

ПОЛУЧЕНИЕ СИГНАЛОВ АМ

Сигналы АМ так же, как и АМ-ПН, можно получать с помощью модулятора с прерывателем и модуляторов, в которых используются нелинейные элементы. В модуляторе с прерывателем (рис. 3.10*а*) сумма модулирующего сигнала $f(t)$ и несущего колебания поступает к прерывателю, который колеблется с частотой ω_c. Действие прерывателя эквивалентно умножению входного сигнала на прямоугольное колебание $p(t)$ частоты ω_c. Спектр прерывис-

того сигнала $u(t)$ можно получить сверткой спектра $[f(t)+$ $+k\cos\omega_c t]$ со спектром $p(t)$, как показано на рис. 3.11. Свертка дает полезный спектр на частотах $\pm\omega_c$ и добавочные нежелатель-

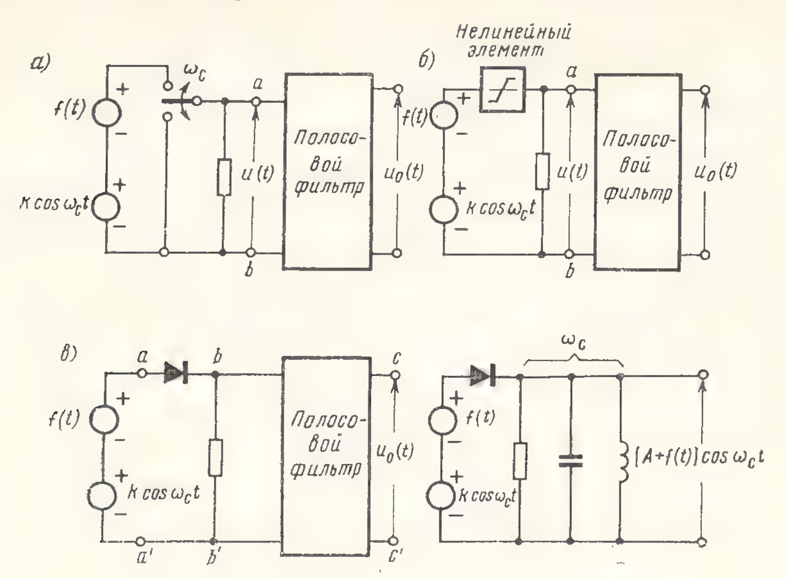


Рис. 3.10. Схемы получения сигналов АМ:
а) модулятор-прерыватель; *б)* модулятор с нелинейным элементом; *в)* модулятор с диодом — прерывателем; *г)* модулятор с нелинейным элементом — диодом

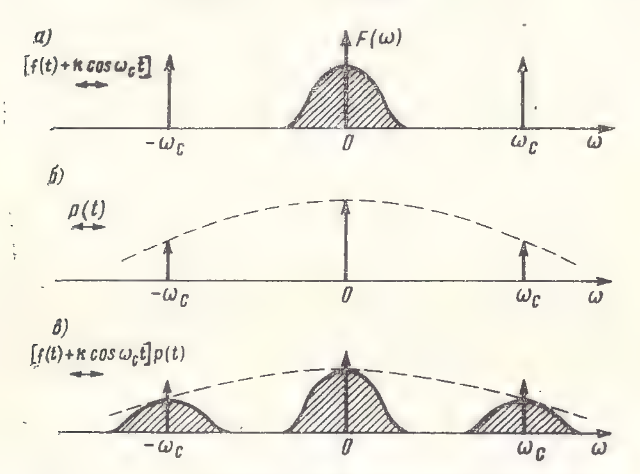


Рис. 3.11. К пояснению действия модулятора-прерывателя:
а) спектр входного сигнала; *б)* спектр периодического сигнала прямоугольной формы; *в)* спектр сигнала на выходе прерывателя

ные спектры на частотах $\omega=0$, $\pm 3\,\omega_c$, $\pm 5\,\omega_c$ и т. д., которые могут быть отфильтрованы полосовым фильтром, настроенным на частоту ω_c. Этот результат нетрудно получить аналитически.

Роль прерывателя может выполнять диод, как показано на рис. 3.10*в*. Если диод идеальный (прямая проводимость бесконечная, обратная — нулевая) и амплитуда несущего колебания превосходит наибольшее значение $f(t)$, то диод действует как ключ, который замыкается при положительных полупериодах несущей и размыкается при отрицательных. Следовательно, диод прерывает входной сигнал с частотой ω_c. Спектр выходного сигнала показан на рис. 3.11*в*. Если такой сигнал пропустить через полосовой фильтр, настроенный на ω_c, то на его выходе получается требуемый АМ сигнал. Заметим, что здесь диод отсекает отрицательную полуволну колебания $[f(t)+k\cos\omega_c t]$. По существу, это однополупериодное выпрямление сигнала. Поэтому такой модулятор известен также как модулятор-выпрямитель.

Модуляцию можно осуществить также с помощью нелинейного элемента (рис. 3.10*б*). Если предположить, что составной элемент, образованный нелинейным элементом и резистором R, имеет вольтамперную характеристику, определяемую ф-лой (3.10), то легко можно показать, что сигнал $u(t)$ состоит из полезного модулированного колебания и нежелательных колебаний, которые отфильтровываются полосовым фильтром, настроенным на ω_c.

ДЕТЕКТИРОВАНИЕ СИГНАЛОВ АМ

Сигналы АМ можно детектировать, используя синхронное детектирование, которое рассматривалось в § 3.2. Возможно, однако, детектирование сигналов АМ значительно более простыми способами с помощью детектора-выпрямителя и детектора огибающей. Эти два вида детекторов на первый взгляд могут показаться одинаковыми, но их действие основано на разных принципах. Детектор-выпрямитель работает по принципу синхронного детектирования[1]), тогда как детектор огибающей является нелинейной цепью, выходное напряжение которой стремится следовать за огибающей входного сигнала. Рассмотрим каждый из этих детекторов.

Детектор-выпрямитель (рис. 3.12), по существу, представляет собой модулятор-выпрямитель, к которому не подводится дополнительное несущее колебание. Он просто выпрямляет модулированный сигнал. Выпрямленный сигнал отличается от входного тем, что отрицательные полупериоды в нем отсутствуют. Это эквива-

---

[1]) Синхронный детектор представляет собой цепь с периодически изменяющимся коэффициентом передачи, причем эти изменения происходят независимо от входного сигнала. В детекторе-выпрямителе подобный эффект вызывается самим входным сигналом. Это делает детектор-выпрямитель нечувствительным к изменениям частоты и фазы входного сигнала. В синхронном детекторе выходное напряжение зависит от разности частот и фаз входного сигнала и опорного напряжения, управляющего проводимостью цепи. — *Прим. ред.*

лентно умножению сигнала на единицу для его положительных значений и на нуль — для отрицательных, т. е. на прямоугольное колебание частоты ω_c Поэтому спектр выпрямленного сигнала получается сверткой спектров модулированного сигнала и колебания $p(t)$. Результат графической свертки показан на рис. 3.13*в*. Из

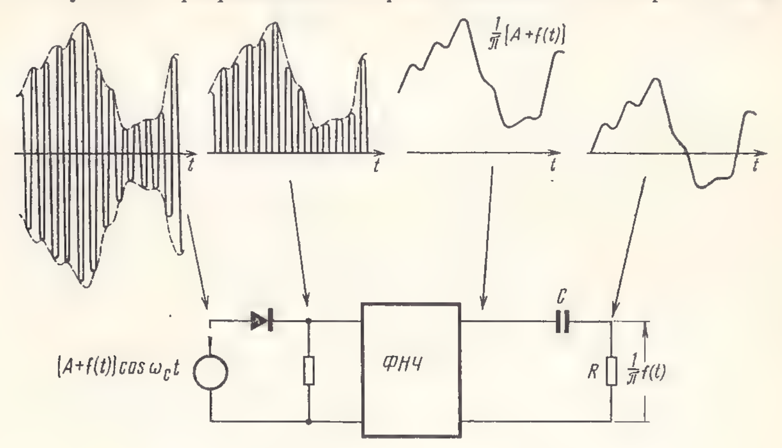



Рис. 3.12. Детектор-выпрямитель сигналов АМ

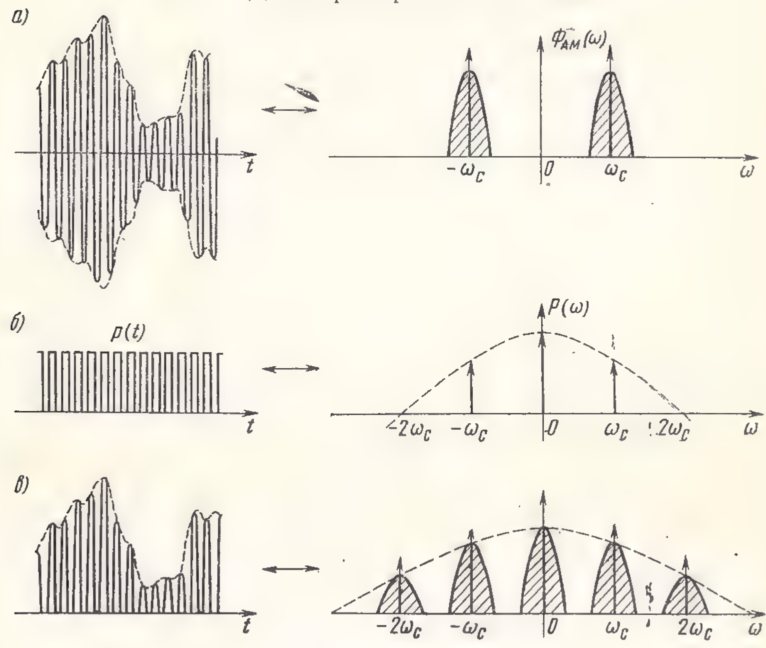


Рис. 3.13. К пояснению действия детектора-выпрямителя:
а) входной сигнал АМ и его спектр; *б)* периодическое колебание прямоугольной формы и его спектр; *в)* выпрямленный сигнал и его спектр

графика видно, что сигнал $f(t)$ можно выделить, пропустив выпрямленный сигнал через фильтр нижних частот. В выходном сигнале этого фильтра содержится постоянная составляющая (единичный импульс в начале координат). Ее можно устранить, включив в схему конденсатор C (рис. 3.12).

Результат свертки легко получить аналитически, зная спектры $\varphi_{AM}(t)$ [ф-ла (3.13)] и $p(t)$ [ф-ла (3.8)]. С другой стороны, можно сразу использовать ф-лу (3.9):

$$\varphi_{AM}(t)\,p(t) \longleftrightarrow \frac{1}{2}\Phi_{AM}(\omega)+\frac{1}{\pi}\sum_{n=-\infty}^{\infty}\frac{(-1)^{(n-1)/2}}{n}\Phi_{AM}(\omega-n\,\omega_c). \quad (3.15)$$

Из ф-лы (3.13) имеем

$$\Phi_{AM}(\omega)=\frac{1}{2}[F(\omega+\omega_c)+F(\omega-\omega_c)+\pi A\,[\delta(\omega+\omega_c)+\delta(\omega-\omega_c)]. \quad (3.16)$$

Подставив ф-лу (3.16) в (3.15), получим полный спектр, показанный на рис. 3.13*в*. Интерес представляет лишь низкочастотная составляющая этого спектра (спектр вблизи $\omega=0$), которая соответствует слагаемому при $n=\pm1$ в ф-ле (3.15). Легко видеть, что выходной сигнал

$$\left.\begin{aligned} u_0(t) &\longleftrightarrow \frac{1}{\pi}F(\omega)+2A\,\delta(\omega) \\ u_o(t) &= \frac{1}{\pi}[A+f(t)]. \end{aligned}\right\}. \quad (3.17)$$

Можно увеличить в два раза выходное напряжение, если вместо однополупериодного использовать двухполупериодный выпрямитель.

Заметим, что выпрямление эквивалентно умножению модулированного сигнала на периодическое колебание (прямоугольной формы) частоты ω_c. Однако необходимо помнить, что умножение осуществляется без дополнительного несущего колебания. Это результат того, что сам модулированный сигнал содержит несущую большого уровня. Если бы несущая в сигнале отсутствовала (как в случае сигнала с подавленной несущей), то выпрямление не было бы эквивалентно умножению входного сигнала на $p(t)$.

В общем случае сигнала АМ, если условие (3.14) $[A+f(t)]>0$ для всех t не выполняется, детектор-выпрямитель применять нельзя. Если же это условие выполняется, то сигнал $[A+f(t)]\cos\omega_c t$ принимает нулевые значения периодически через половину периода несущего колебания и процесс выпрямления эквивалентен умножению сигнала на $p(t)$. Однако если условие $[A+f(t)]>0$ выполняется не для всех значений t, то амплитуда $[A+f(t)]$ изменяет знак с положительного на отрицательный и наоборот. При этом в модулированном сигнале появляются дополнительные переходы че-

рез нулевой уровень, которые не обязательно являются периодическими. Следовательно, выпрямление при этих условиях не эквивалентно умножению сигнала на периодическое прямоугольное колебание $p(t)$ и не может быть использовано для детектирования таких сигналов. Во всех случаях, когда указанное выше условие не выполняется, сигнал $f(t)$ можно выделить синхронным детектированием, для осуществления которого требуется внешнее несущее колебание на приемной стороне.

Из изложенного вытекает также и другая возможность детектирования сигналов с подавленной несущей [и вообще сигналов, не удовлетворяющих условию (3.14)]. Можно добавить к таким сигналам несущее колебание достаточно большой амплитуды, чтобы условие $[A+f(t)]>0$ выполнялось для всех t, а затем для выделения $f(t)$ подвергнуть суммарный сигнал выпрямлению и фильтрации. Этот способ подробно рассматривается в § 3.6.

Детектор огибающей (рис. 3.14*а*) представляет собой выпрямитель с конденсатором, подключенным к его выходным зажимам. В положительный полупериод входного сигнала конденсатор C (рис. 3.14*б*) заряжается до амплитудного значения сигнала. До

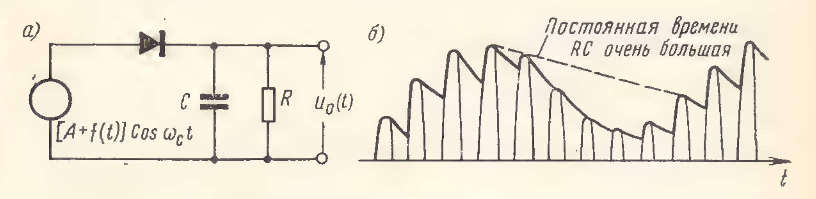


Рис. 3.14. Детектор огибающей:
а) схема детектора; *б)* выходной сигнал при различных значениях RC

тех пор, пока входной сигнал меньше этого значения, диод заперт, так как напряжение на конденсаторе больше, чем значение входного сигнала. Конденсатор сравнительно медленно разряжается через резистор R. Во время следующего положительного полупериода, когда значение входного сигнала близко к амплитудному, оно оказывается больше, чем напряжение на конденсаторе, диод открывается и конденсатор вновь заряжается до амплитудного значения этого нового полупериода.

Таким образом, в течение положительного полупериода конденсатор заряжается до амплитудного значения входного сигнала, причем заряд его почти сохраняется до следующего положительного полупериода. Постоянная времени RC выходного фильтра подбирается так, чтобы экспоненциально убывающее напряжение на конденсаторе в течение периода разряда почти следовало за огибающей (см. задачу 3.9). При этом напряжение на конденсаторе имеет небольшие колебания с частотой ω_c, которые можно отфильтровать еще одним фильтром нижних частот.

Из проведенного анализа следует, что выходное напряжение детектора огибающей в π раз больше, чем в детекторе-выпрямителе [ф-ла (3.17)]. Следовательно, детектор огибающей не только проще детектора-выпрямителя, но и эффективнее его. Поэтому в подавляющем большинстве случаев для детектирования сигналов АМ используется детектор огибающей. Все радиовещательные приемники сигналов АМ имеют детектор огибающей.

МОЩНОСТЬ БОКОВЫХ СОСТАВЛЯЮЩИХ И НЕСУЩЕГО КОЛЕБАНИЯ СИГНАЛА АМ

Само по себе несущее колебание в сигнале АМ не переносит информацию, и с этой точки зрения мощность, передаваемая на несущей частоте, расходуется бесполезно. Интересно определить относительное значение мощности несущего колебания и боковых составляющих (которые переносят информацию). Модулированный сигнал выражается суммой несущего колебания $A\cos\omega_c t$ и боковых составляющих $f(t)\cos\omega_c t$ [ф-ла (3.12)].

Мощность несущего колебания P_c есть среднее квадратическое значение $A\cos\omega_c t$: $P_c = A^2/2$. Мощность боквых составляющих P_s равна среднему квадратическому значению $f(t)\cos\omega_c t$ и составляет половину среднего квадратического значения функции $f(t)$ [см. пример 2.1, ф-ла (2.25)]: $P_s = \frac{1}{2}\overline{f^2(t)}$.

Общая мощность P_t равна сумме P_c и P_s:

$$P_t = P_c + P_s = \frac{1}{2}[A^2 + \overline{f^2(t)}].$$

Процентное содержание мощности боковых составляющих

$$\eta = \frac{P_s}{P_t}100\% = \frac{\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}100\%. \qquad (3.18)$$

Заметим, что при АМ $|f(t)|_{\text{макс}} \leqslant A$. В частном случае, когда $f(t)$ — гармоническая функция, $f(t) = mA\cos\omega_m t$, где m — коэф-

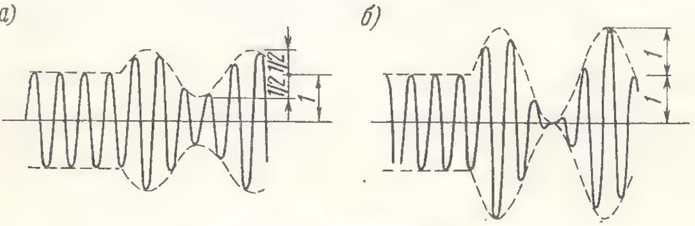


Рис. 3.15. Сигналы АМ:
а) **при $m=0{,}5$;** *б)* **при $m=1$**

фициент модуляции ($m \leqslant 1$). Сигналы АМ при $m=0{,}5$ и $m=1$ показаны на рис. 3.15. В этом случае $\overline{f^2(t)} = (mA)^2/2$ и

$$\eta = \frac{m^2}{2 + m^2}100\%. \qquad (3.19)$$

При $m=1$ значение η составляет $\eta_{макс}=1/3\times 100\%=33{,}3\%$.

Следовательно, при наибольшем значении коэффициента модуляции эффективность передачи равна 33%; 67% мощности переносится несущим колебанием и расходуется бесполезно. При $m<1$ эффективность менее 33%. Заметим, что при АМ-ПН несущая отсутствует и эффективность равна 100%.

3.4. Однополосная амплитудная модуляция

ВВОДНЫЕ ЗАМЕЧАНИЯ

При амплитудной модуляции исходный спектр $F(\omega)$ сдвигается на частоты $\pm\omega_c$, как показано на рис. 3.16*б*. Немодулированный сигнал занимает полосу ω_m (рис. 3.16*а*), тогда как после модуляции полоса сигнала становится $2\omega_m$. Таким образом, рассматри-

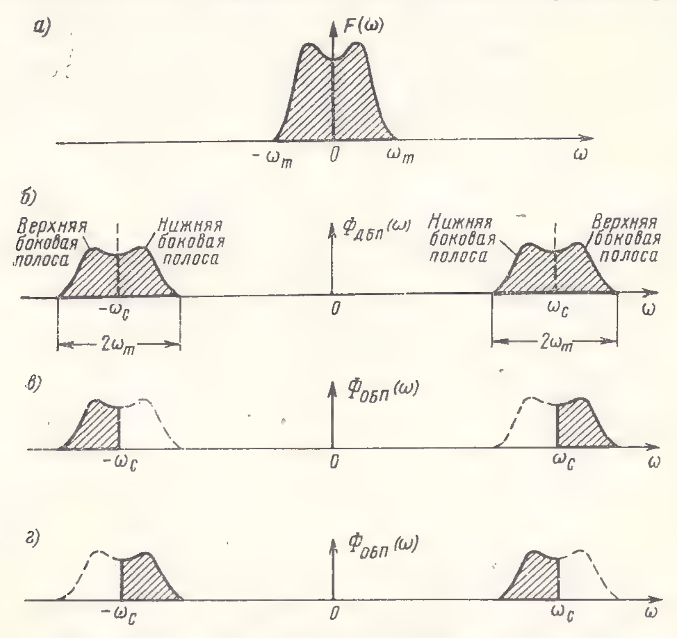


Рис. 3.16. Спектры сигналов двухполосной и однополосной модуляции:

а) спектр модулирующего сообщения $f(t)$; *б)* спектр сигнала ДБП-ПН; *в)* спектр сигнала ОБП-ПН с подавленной нижней боковой полосой; *г)* спектр сигнала ОБП-ПН с подавленной верхней боковой полосой

ваемый до сих пор способ преобразования частоты[1]) характеризуется удвоением ширины спектра сигнала. При этом передается

[1]) Этот способ передачи называется двухполюсным (ДБП).

избыточная информация. Спектр $F(\omega)$ сдвигается на частоты ω_c и $-\omega_c$, в результате чего образуются идентичные спектры, каждый из которых содержит полную информацию относительно $F(\omega)$. Поэтому возникает вопрос: почему бы не передавать только один из них? Однако это невозможно, так как для любого реального сигнала спектр является четной функцией частоты (см. гл. 1). Спектр, который несимметричен относительно вертикальной оси, проходящей через начало координат, не соответствует действительному сигналу и, следовательно, не может быть передан.

Однако существует и другой способ передачи. Как видно из рис. 3.16*б*, спектр вблизи ω_c состоит из двух частей: одна часть расположена выше ω_c и называется верхней боковой полосой частот, другая лежит ниже ω_c и называется нижней боковой полосой. Подобным же образом и спектр вблизи $-\omega_c$ состоит из верхней и нижней боковых полос. Две верхние боковые полосы спектра (или две нижние) содержат всю информацию относительно $F(\omega)$. Следовательно, вместо полного спектра, показанного на рис. 3.16*б*, достаточно передать только либо верхние, либо нижние боковые полосы (показанные на рис. 3.16*в* и *г*). Заметим, что пары боковых полос являются четными функциями частоты и, значит, соответствуют действительному сигналу. Преобразуя частоту однополосного сигнала, можно восстановить исходный сигнал $f(t)$. Для передачи боковых полос в этом случае требуется только половина прежней полосы частот (ω_m вместо $2\omega_m$). Такой способ передачи называется однополосной амплитудной модуляцией с подавленной несущей (ОБП-ПН).

ПОЛУЧЕНИЕ ОДНОПОЛОСНЫХ СИГНАЛОВ[1])

Для получения сигнала ОБП-ПН требуется лишь выделить боковые полосы из модулированного сигнала на выходе балансного модулятора. Сигнал АМ с подавленной несущей, полученный в балансном модуляторе, подводится к соответствующему полосовому фильтру, который пропускает нужную боковую полосу и подавляет все остальные. Этот фильтр должен иметь характеристику, близкую к идеальной прямоугольной форме около частоты ω_c. Иными словами, фильтр должен иметь резкий спад частотной характеристики на частоте ω_c, чтобы не пропускать все составляющие, лежащие по одну сторону от ω_c. Практически легче создать фильтр с резкой отсечкой на более низких частотах. Поэтому, спектр $F(\omega)$ переносится сначала на сравнительно низкую промежуточную частоту $\pm\omega_{c1}$, где одна из боковых полос отфильтровывается. Однополосный спектр на частотах $\pm\omega_{c1}$ из-за несовершенства фильтра-

[1]) Более полные сведения об однополосной амплитудной модуляции приведены в книге М. В. Верзунова, И. В. Лобанова, А. М. Семенова «Однополосная модуляция». Связьиздат, 1962, а также в «Proc. IRE», v. 44, № 12, dec. 1956. — *Прим. ред.*

ции содержит еще остатки нежелательных боковых полос. Затем этот спектр переносится на промежуточную частоту ω_{c2}, где он вновь подвергается фильтрации для удаления остатков ненужной боковой полосы. И только после этого спектр переносится на требуемую несущую частоту ω_c.

Задача фильтрации существенно упрощается, если модулирующий сигнал не содержит значительных низкочастотных составляющих. В этом случае частотные характеристики фильтров, выделяющих однополосный спектр, могут не иметь резкого спада, так как вблизи несущей частоты мощность отфильтровываемых составляющих незначительна. Примером такого сигнала может служить речевой сигнал, в котором мощность низкочастотных составляющих сравнительно невелика.

МЕТОД ФАЗОВОГО СДВИГА

Сигналы ОБП-ПН можно получить методом фазового сдвига (косвенным). Рассмотрим вначале случай гармонического сигнала $f(t)=\cos\omega_s(t)$. Его спектр $F(\omega)$ представляется двумя δ-импульсами на частотах $\pm\omega_s$ (рис. 3.17*а*). Модулированный сигнал равен $\cos\omega_s t\cos\omega_c t$ и имеет спектр, равный $F(\omega)$, но сдвинутый на $\pm\omega_c$ (рис. 3.17*б*). Спектр однополосного сигнала при передаче нижней боковой полосы состоит из двух единичных импульсов на частотах $\pm(\omega_c-\omega_s)$, как показано на рис. 3.17*в*. Очевидно, этому спектру соответствует сигнал $\cos(\omega_c-\omega_s)t$. Следовательно, в частном случае $f(t)=\cos\omega_s t$ получение однополосного сигнала эквивалентно генерированию сигнала $\cos(\omega_c-\omega_s)t$.

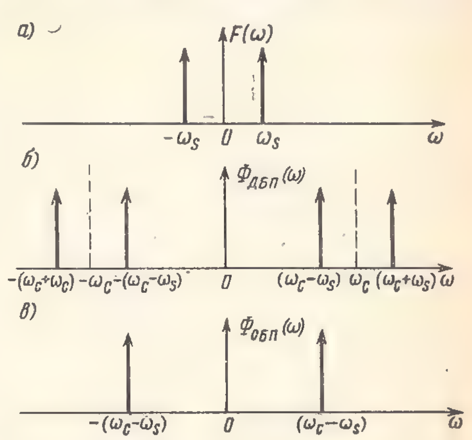


Рис. 3.17. К получению сигналов ОБП-ПН по методу фазового сдвига: *а)* спектр гармонического модулирующего сигнала; *б)* спектр сигнала ДБП-ПН; *в)* спектр сигнала ОБП-ПН

Из тригонометрического тождества имеем

$$\cos(\omega_c-\omega_s)t=\cos\omega_s t\cos\omega_c t+\sin\omega_s t\sin\omega_c t.$$

Таким образом, однополосный сигнал можно получить, суммируя колебания $\cos\omega_s t\cos\omega_c t$ и $\sin\omega_s t\sin\omega_c t$. Колебание $\cos\omega_s t\times\times\cos\omega_c t$ создается с помощью балансного модулятора. Колебание $\sin\omega_s t\sin\omega_c t$ можно выразить как $\cos\left(\omega_s t-\frac{\pi}{2}\right)\cos\left(\omega_c t-\frac{\pi}{2}\right)$. Следовательно, и этот сигнал можно получить в балансном модулято-

ре при условии, что сигналы $\cos\omega_s t$ и $\cos\omega_c t$ сдвигаются по фазе на $-\pi/2$ (рис. 3.18). Хотя этот результат найден для частного случая $f(t)=\cos\omega_s t$, он справедлив для любого сигнала. Напом-

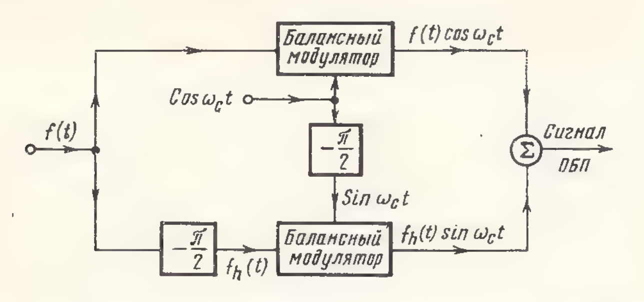


Рис. 3.18. Структурная схема модулятора сигналов ОБП-ПН

ним, что любое колебание можно выразить как непрерывную сумму гармонических сигналов. Следовательно, сигнал ОБП с подавленной несущей в случае произвольной модулирующей функции $f(t)$ записывается в виде

$$\varphi_{ОБП}(t) = f(t)\cos\omega_c t + f_h(t)\sin\omega_c t, \tag{3.20}$$

где $f_h(t)$ — колебание, получаемое из $f(t)$ сдвигом фаз всех частотных составляющих на $-\pi/2$.

Докажем справедливость этого результата для произвольного сигнала $f(t)$. Фазосдвигающая цепь (фазовращатель), которая изменяет фазу частотных составляющих на $-\pi/2$, имеет единичный коэффициент передачи. Таким образом, амплитуды частотных составляющих остаются неизменными, но их фазы на положительных частотах сдвигаются на $-\pi/2$. Поскольку фазовый спектр является нечетной функцией частоты, фазы всех составляющих на отрицательных частотах сдвигаются на $+\pi/2$. Амплитудно-частотная и фазо-частотная характеристики фазосдвигающей цепи показаны на рис. 3.19:

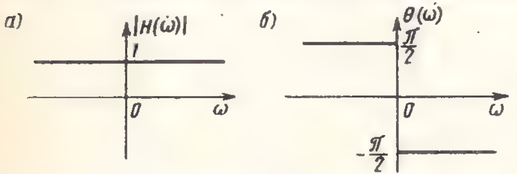


Рис. 3.19. Передаточная функция фазосдвигающей цепи:
а) амплитудно-частотная характеристика;
б) фазо-частотная характеристика

$$|H(\omega)| = 1, \quad \theta(\omega) = \frac{\pi}{2} - \pi u_1(\omega).$$

Следовательно, передаточная функция $H(\omega)$ такой фазосдвигающей цепи записывается как

$$H(\omega) = |H(\omega)|\, e^{i\theta(\omega)} = e^{i\left[\frac{\pi}{2} - \pi u_1(\omega)\right]} = i e^{-i\pi u_1(\omega)}. \tag{3.21}$$

Если $f(t) \longleftrightarrow F(\omega)$, то

$$f_h(t) \longleftrightarrow \mathrm{i}\, F(\omega)\, \mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega)}\,. \qquad (3.22)$$

Слагаемое $f(t)\cos\omega_c t$ в ф-ле (3.20) имеет преобразование Фурье, определяемое ф-лой (3.1).

Преобразование Фурье $f_h(t)\sin\omega_c t$ на основании ф-л (3.23) и (1.116в) можно записать в виде:

$$f_h(t)\sin\omega_c t \longleftrightarrow -\frac{1}{2}\left[F(\omega+\omega_c)\,\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega+\omega_c)} - F(\omega-\omega_c)\,\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega-\omega_c)}\right], \qquad (3.23)$$

$$[f(t)\cos\omega_c t + f_h(t)\sin\omega_c t] \longleftrightarrow \frac{1}{2}F(\omega-\omega_c)\left[1+\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega-\omega_c)}\right] + \frac{1}{2}F(\omega+\omega_c)\left[1-\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega+\omega_c)}\right]. \qquad (3.24)$$

Заметим, что $u_1(\omega-\omega_c) = \begin{cases} 1 & \text{при } \omega > \omega_c, \\ 0 & \text{при } \omega < \omega_c. \end{cases}$

Следовательно,

$$1+\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega-\omega_c)} = \begin{cases} 2 & \text{при } \omega < \omega_c, \\ 0 & \text{при } \omega > \omega_c. \end{cases}$$

Но по определению это есть $2u_1(\omega_c-\omega)$.

$$1+\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega-\omega_c)} = 2u_1(\omega_c-\omega)\,. \qquad (3.25а)$$

Подобным же образом

$$1-\mathrm{e}^{-\mathrm{i}\,\pi u_1(\omega+\omega_c)} = 2u_1(\omega_c+\omega). \qquad (3.25б)$$

Подставив эти соотношения в ф-лу (3.24), получим

$$[f(t)\cos\omega_c t + f_h(t)\sin\omega_c t] \longleftrightarrow [F(\omega-\omega_c)\,u_1(\omega_c-\omega) + F(\omega+\omega_c)\,u_1(\omega+\omega_c)]. \qquad (3.26)$$

Правая часть ф-лы (3.26) выражает спектр нижних боковых полос $[F(\omega-\omega_c)+F(\omega+\omega_c)]$. Слагаемое $F(\omega-\omega_c)u_1(\omega_c-\omega)$ представляет нижнюю боковую полосу спектра $F(\omega-\omega_c)$, так как $u_1(\omega_c-\omega)=0$ при $\omega>\omega_c$. Подобным же образом $F(\omega+\omega_c)u_1(\omega+\omega_c)$ представляет нижнюю боковую полосу спектра $F(\omega+\omega_c)$, так как $u_1(\omega+\omega_c)=0$ при $\omega<-\omega_c$. Таким образом, сигнал, выражаемый ф-лой (3.24), действительно является однополосным с подавленными верхними боковыми полосами. Можно показать, что если суммирование заменить вычитанием $f_h(t)\sin\omega_c t$ из $f(t)\cos\omega_c t$, то в результате получится однополосный сигнал с подавленными нижними боковыми полосами. Таким образом, в общем случае сигнал ОБП-ПН $\varphi_{\text{ОБП}}(t)$ можно записать как

$$\varphi_{\text{ОБП}}(t) = f(t)\cos\omega_c t \pm f_h(t)\sin\omega_c t, \qquad (3.27)$$

где знак плюс соответствует однополосному сигналу с нижними боковыми полосами, а минус — с верхними.

Колебание $f_h(t)$ есть отклик фазосдвигающей цепи (рис. 3.19)

на сигнал $f(t)$. Функцию $f_h(t)$ можно выразить через $f(t)$, используя ф-лу (3.22).

Заметим, что

$$e^{-i\pi u_1(\omega)} = \begin{cases} -1 \text{ при } \omega > 0, \\ 1 \text{ при } \omega < 0 \end{cases} = -\operatorname{sign}(\omega).$$

Следовательно,

$$f_h(t) \longleftrightarrow -i\, F(\omega) \operatorname{sign}(\omega). \tag{3.28}$$

В соответствии с ф-лой (3.28) имеем

$$\frac{i}{\pi t} \longleftrightarrow \operatorname{sign}(\omega).$$

Применяя к ф-ле (3.28) теорему временно́й свертки, получаем [1])

$$f_h(t) = \frac{1}{\pi} f(t) * \frac{1}{t} = \frac{1}{\pi} \int\limits_{-\infty}^{\infty} \frac{f(\tau)}{t-\tau}\, d\tau. \tag{3.29}$$

ДЕТЕКТИРОВАНИЕ СИГНАЛОВ ОБП-ПН

Для выделения сообщения $f(t)$ из однополосного сигнала необходимо сдвинуть спектр, показанный на рис. 3.16*в* или *г*, в первоначальное положение (на частоту $\omega=0$). Это легко выполняется при синхронном детектировании. Умножение однополосного сигнала на $\cos\omega_c t$ (синхронное детектирование) эквивалентно свертке спектра однополосного сигнала со спектром колебания $\cos\omega_c t$ (состоящим из двух единичных импульсов на частотах $\pm\omega_c$). Эта свертка показана на рис. 3.20 для однополосного сигнала с верхней боковой полосой. В результате свертки получается спектр $F(\omega)$ и дополнительный однополосный сигнал с несущей $2\omega_c$. Последний может быть устранен фильтром нижних частот.

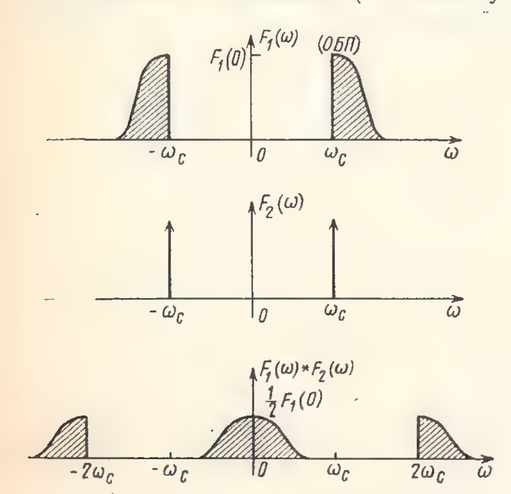


Рис. 3.20. К пояснению детектирования сигналов ОБП-ПН

[1]) Функция $f_h(t)$ есть преобразование Гильберта сигнала $f(t)$. Она называется квадратурной для $f(t)$, так как каждая частотная составляющая $f(t)$ находится в фазовой квадратуре с составляющей $f_h(t)$, как видно из рис. 3.19. Интеграл в ф-ле (3.29) является несобственным, и при $t=\tau$ берется главное значение интеграла по Коши.

Этот результат можно получить аналитически. Выходное напряжение синхронного детектора $u_d(t)$ [см. ф-лу (3.28)]:

$$u_d(t) = \varphi_{ОБП}(t)\cos\omega_c t = f(t)\cos^2\omega_c t \pm f_h(t)\sin\omega_c t\cos\omega_c t =$$

$$= \frac{1}{2} f(t) + [f(t)\cos 2\omega_c t \pm f_h(t)\sin 2\omega_c t]. \quad (3.30)$$

Первое слагаемое этой суммы представляет собой сообщение, а слагаемое в квадратных скобках — однополосный сигнал с несущей $2\omega_c$. Следовательно, при синхронном детектировании получается исходное сообщение $f(t)$ и однополосный сигнал с несущей $2\omega_c$, который затем отфильтровывается.

Синхронное детектирование сигналов ОБП может выполняться любой из схем, показанных на рис. 3.7. Для синхронного детектирования необходимо генерировать местное несущее колебание, имеющее частоту ω_c и находящееся в фазе с несущей принимаемого сигнала. Любые ошибки в частоте и фазе местной несущей приводят к искажениям сообщения на выходе детектора. Природа этих искажений подробно рассматривается в § 3.5.

ДЕТЕКТИРОВАНИЕ СИГНАЛА ОБП С БОЛЬШИМ УРОВНЕМ НЕСУЩЕЙ

Рассмотрим сигнал ОБП с большим уровнем несущей. Такой сигнал можно представить в виде

$$\varphi(t) = A\cos\omega_c t + [f(t)\cos\omega_c t \pm f_h(t)\sin\omega_c t].$$

Следовательно, сообщение $f(t)$ можно выделить из $\varphi(t)$ синхронным детектором (умножая $\varphi(t)$ на $\cos\omega_c t$). Однако если амплитуда A несущей достаточно велика, то $f(t)$ можно выделить из $\varphi(t)$ с помощью детектора-выпрямителя. Это легко показать, переписав $\varphi(t)$ следующим образом:

$$\varphi(t) = [A + f(t)]\cos\omega_c t + f_h(t)\sin\omega_c t = U(t)\cos(\omega_c t + \theta),$$

где $U(t) = \{[A + f(t)]^2 + f_h^2(t)\}^{1/2}$, $\theta(t) = \operatorname{arc\,tg}\{f_h(t)/[A + f(t)]\}$.

Очевидно, $U(t)$ есть огибающая однополосного сигнала $\varphi(t)$. Если сигнал $\varphi(t)$ подается на вход детектора огибающей, то на его выходе появится сигнал

$$U(t) = \{[A + f(t)]^2 + f_h^2(t)\}^{1/2} =$$

$$= A[1 + 2f(t)/A + f^2(t)/A^2 + f_h^2(t)/A^2]^{1/2}.$$

Если $A \gg |f(t)|$, то $A \gg |f_h(t)|$, и слагаемыми $f^2(t)/A^2$ и $f_h^2(t)/A^2$ можно пренебречь. В этом случае

$$U(t) \approx A[1 + 2f(t)/A]^{1/2}.$$

Используя биномиальное разложение и отбрасывая слагаемые высшего порядка (так как $f(t)/A \ll A$), получаем

$$U(t) \approx A[1 + f(t)/A] = A + f(t).$$

Таким образом, при большом уровне несущей огибающая сигнала $\varphi(t)$ имеет форму сообщения $f(t)$, и возможно детектирование с помощью детектора огибающей. Передача сигнала ОБП с большим уровнем несущей используется в телевизионном вещании[1]).

3.5. Влияние частотных и фазовых ошибок опорного колебания на синхронное детектирование

ВВОДНЫЕ ЗАМЕЧАНИЯ

Сигналы АМ можно детектировать с помощью детектора-выпрямителя или детектора огибающей без генерации местного несущего колебания в приемнике, так как передаваемый сигнал содержит несущее колебание большой амплитуды. Однако в системах с подавленной несущей (ДБП-ПН и ОБП-ПН) необходимо генерировать местную несущую в приемнике для выполнения синхронного детектирования, причем частота и фаза местной несущей должны быть точно равны частоте и фазе несущей передатчика. Любые отклонения частоты и фазы местной несущей приводят к появлению искажений сообщения на выходе детектора. Выше отмечалось влияние частотной ошибки [ф-ла (3.4)]. Теперь этот вопрос будет рассмотрен более подробно для сигналов ДБП и ОБП.

СИГНАЛЫ ДБП-ПН

Пусть принимаемый сигнал есть $f(t)\cos\omega_c t$, а местная несущая — $\cos[(\omega_c+\Delta\omega)t+\varphi]$. Частотная и фазовая ошибки местной несущей обозначены $\Delta\omega$ и φ соответственно. Синхронное детектирование сводится к умножению принимаемого сигнала на местное несущее колебание и пропусканию полученного произведения через фильтр нижних частот, как показано на рис. 3.1*е*. Произведение принимаемого сигнала и местного несущего колебания записывается в виде

$$u_d(t)=f(t)\cos\omega_c t\cos[(\omega_c+\Delta\omega)t+\varphi]=$$

$$=\frac{1}{2}f(t)\{\cos[\Delta\omega\, t+\varphi]+\cos[(2\omega_c+\Delta\omega)\,t+\varphi]\}. \tag{3.31}$$

Второе слагаемое в правой части ф-лы (3.31) представляет собой колебание, спектр которого расположен вблизи частоты $(2\omega_c+ +\Delta\omega)$. Оно отфильтровывается фильтром нижних частот с частотой среза ω_m (рис. 3.1). Напряжение на выходе фильтра

$$u_o(t)=\frac{1}{2}f(t)\cos[\Delta\omega t+\varphi]. \tag{3.32}$$

Из этого выражения следует, что выходное напряжение отли-

---

[1]) В телевизионных системах используется несколько видоизмененная форма однополосной передачи — так называемая передача с частично-подавленной боковой полосой (см. § 3.8).

чается от $f(t)$ и, таким образом, является искаженным. Заметим, что если $\Delta\omega$ и φ равны нулю (фазовая и частотная ошибки отсутствуют), то $u_o(t)=\frac{1}{2}f(t)$, как и следовало ожидать.

Рассмотрим два частных случая:

1) $\Delta\omega=0$, $\varphi\neq0$ (только фазовая ошибка),
2) $\Delta\omega\neq0$, $\varphi=0$ (только частотная ошибка).

Если $\Delta\omega=0$, то ф-ла (3.32) принимает вид

$$u_0(t)=\frac{1}{2}f(t)\cos\varphi.$$

Выходное напряжение детектора пропорционально $f(t)$, если φ — постоянная величина. Выходное напряжение максимально при $\varphi=0$ и минимально (равно нулю) при $\varphi=\pm\pi/2$. Таким образом, фазовая ошибка в местной несущей вызывает ослабление выходного сигнала пропорционально косинусу этой ошибки. Однако, если φ — постоянная величина, то форма сообщения не искажается. На практике же обычно фазовая ошибка φ изменяется во времени случайным образом (в результате изменений пути распространения радиоволн вследствие случайных колебаний в ионосфере). При этом коэффициент передачи детектора $\cos\varphi$ случайно изменяется, что, очевидно, нежелательно.

Если $\varphi=0$, а $\Delta\omega\neq0$, ф-ла (3.32) принимает вид

$$u_o(t)=\frac{1}{2}f(t)\cos\Delta\omega t.$$

Теперь выходное напряжение не является просто ослабленной копией сообщения, оно искажено. Поскольку обычно величина $\Delta\omega$ мала, выходное напряжение детектора есть сообщение $f(t)$, умноженное на гармоническое низкочастотное колебание. Уровень сообщения $f(t)$ периодически изменяется во времени. Это более серьезный вид искажений, поэтому важно, чтобы местный генератор был синхронизирован. Такая синхронизация обычно выполняется с помощью обратной связи. Напряжения, пропорциональные частотной и фазовой ошибкам, необходимые для управления местным генератором по цепи обратной связи, создаются в специаль-

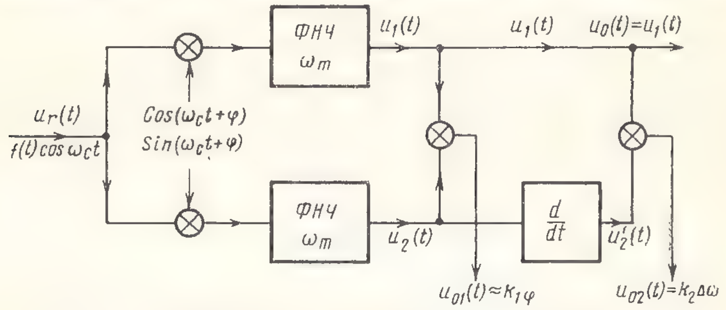


Рис. 3.21. Структурная схема устройства для получения сигналов фазовой и частотной ошибок опорной несущей приемника

ном устройстве (рис. 3.21). Выходное напряжение местного генератора преобразуется в две квадратурные составляющие (косинусоидальную и синусоидальную). Выходное напряжение $u_1(t)$ равно произведению входного сигнала $f(t)\cos\omega_c t$ на $\cos(\omega_c t+\varphi)$, пропущенному через фильтр нижних частот *(ФНЧ)*, который отфильтровывает все составляющие с частотами, выше ω_m:

$$f(t)\cos\omega_c t\cos(\omega_c t+\varphi)=\frac{1}{2}f(t)\cos\varphi+\frac{1}{2}f(t)\cos(2\omega_c t+\varphi).$$

Следовательно, $u_1(t)=\frac{1}{2}f(t)\cos\varphi$.

Подобным же образом $u_2(t)=\frac{1}{2}f(t)\sin\varphi$.

Произведение этих сигналов

$$u_1(t)\,u_2(t)=\frac{1}{4}f^2(t)\cos\varphi\sin\varphi=\frac{1}{8}f^2(t)\sin 2\varphi\approx\frac{1}{4}f^2(t)\,\varphi\ (\text{при }\varphi\ll 1). \tag{3.33}$$

Таким образом, выходное напряжение $u_{o1}(t)$ пропорционально фазовой ошибке φ. Полярность выходного напряжения зависит от того, положительна или отрицательна ошибка. Полученное напряжение можно использовать в качестве управляющего для регулировки фазы местного генератора. Такая система автоматической регулировки фазы изменяет не только фазу, но также в узких пределах и частоту местного генератора.

Можно показать, что если напряжение $u_2(t)$ продифференцировать, а затем умножить на $u_1(t)$, то произведение $u_1(t)u_2'(t)$ будет пропорционально частотной ошибке. Это обусловлено тем, что изменение мгновенной частоты сигнала равно скорости изменения фазы. Следовательно, частотная ошибка в сигнале $\cos(\omega_c t+\varphi)$ равна $\Delta\omega=d\varphi/dt$. Таким образом,

$$u_2'(t)=\frac{du_2}{dt}=\frac{1}{2}\left[f(t)\cos\varphi\frac{d\varphi}{dt}+\varphi'(t)\sin\varphi\right].$$

При малых значениях φ $\sin\varphi\approx 0$, $\cos\varphi\approx 1$, поэтому

$$u_2'(t)\approx\frac{1}{2}f(t)\frac{d\varphi}{dt}=\frac{1}{2}\Delta\omega f(t).$$

и

$$u_1(t)\,u_2'(t)=\frac{1}{4}f^2(t)\,\Delta\omega\cos\varphi\approx\frac{1}{4}f^2(t)\,\Delta\omega. \tag{3.34}$$

Итак, произведение $u_1(t)u_2'(t)$ пропорционально частотной ошибке $\Delta\omega$. Когда частотная и фазовая ошибки устраняются системой фазовой автоподстройки частоты, достигается синхронизм; при этом $u_1(t)=\frac{1}{2}f(t)$, а $u_2(t)=0$.

Сигнал ОБП на входе приемника[1]) определяется по формуле (3.27). Пусть местная несущая есть $\cos[(\omega_c+\Delta\omega)t+\varphi]$. Произведение входного сигнала и местной несущей

$$u_d(t)=\varphi_{ОБП}(t)\cos[(\omega_c+\Delta\omega)t+\varphi]=$$
$$=[f(t)\cos\omega_c t+f_h(t)\sin\omega_c t]\cos[(\omega_c+\Delta\omega)t+\varphi]=$$
$$=\frac{1}{2}f(t)\{\cos[\Delta\omega t+\varphi]+\cos[(2\omega_c+\Delta\omega)t+\varphi]\}-$$
$$-\frac{1}{2}f_h(t)\{\sin[\Delta\omega t+\varphi]-\sin[(2\omega_c+\Delta\omega)t+\varphi]\}.$$

Слагаемые с частотами $2\omega_c+\Delta\omega$ отфильтровываются фильтром нижних частот, выходное напряжение которого

$$u_o(t)=\frac{1}{2}f(t)\cos[\Delta\omega t+\varphi]-\frac{1}{2}f_h(t)\sin[\Delta\omega t+\varphi]. \qquad (3.35)$$

Заметим, что если $\Delta\omega=0$ и $\varphi=0$, то $u_o(t)=\frac{1}{2}f(t)$, как и следовало ожидать. Интересно сравнить влияние фазовой и частотных ошибок в системах с ОБП и ДБП. Если $\Delta\omega=0$, то, как уже отмечалось, в случае сигнала ДБП сообщение не искажается, хотя и изменяется по величине пропорционально $\cos\varphi$. В случае сигнала ОБП при $\Delta\omega=0$

$$u_o(t)=\frac{1}{2}[f(t)\cos\varphi-f_h(t)\sin\varphi]. \qquad (3.36)$$

Из этого выражения следует, что выходное напряжение содержит нежелательный сигнал $f_h(t)\sin\varphi$, который не может быть отфильтрован. Покажем, что возникающие искажения являются фазовыми. Если все частотные составляющие сообщения $f(t)$ сдвигаются по фазе на φ *рад*, то получается колебание $u_o(t)$ в виде (3.36). Для доказательства этого определим преобразование Фурье сигнала (3.36)

$$U_o(\omega)=\frac{1}{2}[F(\omega)\cos\varphi-F_h(\omega)\sin\varphi]. \qquad (3.37)$$

Но из ф-лы (3.22) следует, что $F_h(\omega)=\mathrm{i}F(\omega)e^{-\mathrm{i}\pi u_1(\omega)}$. Следовательно,

$$U_o(\omega)=\frac{1}{2}F(\omega)\left[\cos\varphi-\mathrm{i}\,e^{-\mathrm{i}\pi u_1(\omega)}\sin\varphi\right]=\begin{cases}\frac{1}{2}F(\omega)e^{\mathrm{i}\varphi} & \text{при } \omega>0,\\ \frac{1}{2}F(\omega)e^{-\mathrm{i}\varphi} & \text{при } \omega<0.\end{cases}$$

---

[1]) Здесь рассматривается лишь сигнал ОБП с нижней боковой полосой. Однако все выводы справедливы и для сигналов с верхней боковой полосой.

Из ф-лы (3.37) следует, что $u_0(t)$ можно получить из $f(t)$, сдвигая фазы всех частотных составляющих на φ *рад* (заметим, что на отрицательных частотах фазовый сдвиг равен $-\varphi$ благодаря свойству антисимметрии фазового спектра). Таким образом, фазовая ошибка местной несущей приводит к появлению фазовых искажений сообщения на выходе детектора. Выходное сообщение есть искаженная форма $f(t)$, в которой каждая частотная составляющая $f(t)$ имеет постоянный фазовый сдвиг. При передаче речи фазовые искажения несущественны, так как ухудшается только ее качество, но разборчивость сохраняется. Однако при передаче музыки и изображений эти искажения недопустимы.

Влияние частотной ошибки в системе ОБП подобно ее влиянию в системе ДБП (при малых значениях $\Delta\omega$). Это легко показать, положив $\varphi=0$ в ф-ле (3.35).

3.6. Детектирование сигналов с подавленной несущей посредством добавления несущего колебания

СИГНАЛЫ ДБП-ПН

Как уже отмечалось, можно детектировать сигналы с подавленной несущей, добавляя несущую к сигналу в приемнике. Этот метод применим как к сигналам ДБП-ПН, так и к сигналам ОБП-ПН. После восстановления несущей достаточного уровня можно использовать либо детектор-выпрямитель, либо детектор огибающей. Для предотвращения искажений фаза и частота вводимой несущей должны быть синхронизированы с несущей передатчика.

Рассмотрим детектирование сигналов ДБП-ПН.

Принимаемый сигнал равен $f(t)\cos\omega_c t$. Пусть восстанавливаемая несущая равна $A\cos(\omega_c t+\varphi)$. Тогда получающийся сигнал $r(t)$ определится выражением

$$r(t)=f(t)\cos\omega_c t+A\cos(\omega_c t+\varphi)= \quad (3.38)$$

$$=[f(t)+A\cos\varphi]\cos\omega_c t-[A\sin\varphi]\sin\omega_c t=$$

$$=\sqrt{[A+f(t)]^2-2Af(t)[1-\cos\varphi]}\cos(\omega_c t+\theta)= \quad (3.39)$$

$$=U(t)\cos(\omega_c t+\theta), \quad (3.40)$$

где

$$U(t)=\{[A+f(t)]^2-2Af(t)[1-\cos\varphi]\}^{1/2}, \quad (3.41)$$

$$\theta=\operatorname{arc\,tg}\{A\sin\varphi/[f(t)+A\cos\varphi]\}.$$

Из ф-лы (3.40) вытекает, что $U(t)$ есть огибающая сигнала $r(t)$. Если $\varphi=0$, то огибающая $U(t)=A+f(t)$.

Следовательно, можно выделить сообщение $f(t)$ из сигнала $r(t)$ с помощью детектора-выпрямителя или детектора огибающей.

Заметим, что в этом случае

$$r(t) = [A + f(t)] \cos \omega_c t.$$

Это колебание есть не что иное, как сигнал АМ, который можно детектировать детектором-выпрямителем или детектором огибающей при условии, что $A + f(t) > 0$ при всех значениях t. Однако, если фазовая ошибка φ не равна нулю, возникают искажения. Формулу (3.41) можно переписать в виде

$$U(t) = A \left\{ 1 + \frac{2f(t)}{A} \cos \varphi + \left[\frac{f(t)}{A} \right]^2 \right\}^{1/2}.$$

Если $A \gg |f(t)|$, то

$$U(t) \approx A + f(t) \cos \varphi. \tag{3.42}$$

Полезной составляющей является $f(t) \cos \varphi$. Если φ — постоянная величина, то полезное сообщение не искажается, но ослабляется пропорционально $\cos \varphi$, как и при синхронном детектировании.

Далее рассмотрим случай частотной ошибки $\Delta\omega \neq 0$ $(\varphi = 0)$.

В этом случае

$$r(t) = f(t) \cos \omega_c t + A \cos [\omega_c t + \Delta\omega\, t].$$

Заметим, что это выражение совпадает с ф-лой (3.38), если в последней φ заменить на $\Delta\omega t$. Следовательно,

$$U(t) \approx A + f(t) \cos \Delta\omega\, t \quad \text{при} \quad A \gg |f(t)|.$$

Эти искажения также подобны тем, которые получались при синхронном детектировании, когда местный генератор имел частотную ошибку.

СИГНАЛЫ ОБП-ПН

Рассмотрим сначала местную несущую только с фазовой ошибкой:

$$r(t) = \varphi_{\text{ОБП}}(t) + A \cos(\omega_c t + \varphi) =$$

$$= [f(t) \cos \omega_c t + f_h(t) \sin \omega_c t] + A \cos(\omega_c t + \varphi) =$$

$$= [A \cos \varphi + f(t)] \cos \omega_c t + [f_h(t) - A \sin \varphi] \sin \omega_c t = U(t) \cos(\omega_c t + \theta),$$

где

$$U(t) = [A^2 + 2Af(t) \cos \varphi - 2Af_h(t) \sin \varphi + f^2(t) + f_h^2(t)]^{1/2} \tag{3.43}$$

и

$$\theta = \operatorname{arc\,tg} \{[A \sin \varphi - f_h(t)] / [A \cos \varphi + f(t)]\}.$$

Если $A \gg |f(t)|$, то $A \gg |f_h(t)|$ и

$$U(t) \approx A + f(t) \cos \varphi - f_h(t) \sin \varphi.$$

На выходе детектора-выпрямителя или детектора огибающей получим $U(t)$. Постоянное напряжение A можно блокировать конденсатором, и получающийся выходной сигнал

$$u_o(t) = f(t) \cos \varphi - f_h(t) \sin \varphi. \tag{3.44}$$

Это выражение с точностью до постоянного множителя совпадает с ф-лой (3.36). Выше было показано (§ 3.5), что колебание $u_o(t)$ в ф-ле (3.44) представляет собой исходное сообщение $f(t)$ с фазовыми искажениями. Таким образом, фазовая ошибка местной несущей приводит к появлению таких же искажений сообщения, как и при синхронном детектировании. Нетрудно показать, что и частотная ошибка в рассматриваемом способе детектирования приводит к тому же результату, что и при синхронном детектировании.

Заметим, что при $\varphi=0$ ф-ла (3.44) принимает вид $u_o(t)=$ $=A+f(t)$.

Итак, если в принимаемый сигнал ОБП-ПН вводится несущая с большим уровнем, детектирование может выполняться детектором-выпрямителем или детектором огибающей.

Несущую не обязательно восстанавливать в приемнике, ее можно вводить и в передатчике. В последнем случае принимаемый сигнал является однополосным с большим уровнем несущей, и его можно детектировать детектором огибающей (или детектором-выпрямителем) подобно сигналу АМ. Однополосный сигнал с большим уровнем несущей можно получить из сигнала АМ, подавляя одну из боковых полос. Такой способ передачи имеет преимущества как перед ОБП, так и перед АМ, поскольку он требует лишь половины полосы, занимаемой сигналом АМ. В то же время детектирование осуществляется простым детектором огибающей, как и в случае сигналов АМ. Такой способ в несколько измененном виде применяется для передачи телевидения (см. § 3.8).

Из изложенного следует, что сигналы с подавленной несущей можно детектировать, вводя несущую с достаточно большим уровнем в приемнике и применяя детектор-выпрямитель или детектор огибающей. Этого следовало ожидать для сигналов ДБП, так как добавление в сигнал ДБП-ПН несущей большого уровня превращает его в сигнал АМ. Результат не столь очевиден для сигналов ОБП-ПН. Его можно объяснить следующим образом. Если в сигнал ОБП-ПН вводится несущая большого уровня, она служит для него как бы подставкой. Поэтому получающийся сигнал имеет пересечения нулевого уровня, весьма близкие к тем, которые дает несущее колебание. Следовательно, выпрямление в этом случае равноценно умножению сигнала на колебание $p(t)$ прямоугольной формы. Очевидно, это эквивалентно синхронному детектированию.

3.7. Сравнение различных систем АМ

Выше были рассмотрены различные системы АМ (ДБП и ОБП) и системы АМ-ПН (ДБП-ПН и ОБП-ПН). Интересно сравнить эти системы с разных точек зрения.

С точки зрения осуществления приемника система АМ имеет преимущество перед АМ-ПН, так как она требует более простого детектора (по этой причине все вещательные системы связи используют АМ), кроме того, сигналы АМ легче получать при высо-

ком уровне мощности по сравнению с сигналами с подавленной несущей. Для получения последних требуются балансные модуляторы, которые несколько труднее реализовать.

Системы с подавленной несущей имеют преимущество перед АМ в том, что они требуют значительно меньшую мощность для передачи той же информации. Это приводит к снижению стоимости передатчика по сравнению с передатчиком АМ. Но приемники в этих системах более сложные, так как они должны генерировать местное несущее колебание, совпадающее по частоте и фазе с несущей входного сигнала. Для систем связи, в которых на один передатчик приходится лишь несколько приемников, это усложнение приемника может быть оправдано.

Однако влияние селективных замираний (благодаря многолучевому распространению) значительно более неблагоприятно сказывается на АМ, чем на АМ-ПН.

Сигналы к приемнику приходят различными путями. Они отражаются ионосферой и земной поверхностью. Свойства ионосферы случайно изменяются во времени, приводя к случайному изменению длины путей. Поэтому фазы сигналов, приходящих к приемнику по разным путям, изменяются случайным образом. Это явление называется замираниями. Замирания зависят и от частоты, и вследствие этого их вредное влияние оказывается еще более серьезным. Таким образом, несущая и каждая из боковых полос подвержены различным замираниям. Замирания этого типа называются *селективными*. При селективных замираниях боковые полосы искажаются по-разному, что приводит к искажениям передаваемого сигнала. Они также изменяют соотношения между уровнями несущего и боковых колебаний. Это может привести к столь значительному ослаблению несущей, что условие (3.14) не будет удовлетворяться. Когда такой сигнал детектируется детектором огибающей (или детектором-выпрямителем), выходное сообщения оказывается сильно искаженным. Влияние селективных замираний более значительно проявляется на высоких частотах. Поэтому на высоких частотах системы с подавленной несущей более предпочтительны.

Далее сравним системы ДБП-ПН и ОБП-ПН. Как уже отмечалось, занимаемая сигналом ОБП полоса частот составляет половину полосы сигнала ДБП[1]). Кроме того, в системах ДБП селективные замирания изменяют фазовые соотношения двух боковых полос, что приводит к искажениям выходного сообщения. В системах ОБП такой возможности нет, поскольку имеется лишь одна боковая полоса. В условиях длинного пути распространения влияние селективных замираний оказывается более вредным для ДБП, чем для ОБП.

По этим причинам системы с ОБП-ПН более предпочтительны, чем с ДБП-ПН. Системы с ОБП-ПН используются на линиях связи большой протяженности, в частности при передаче речи, где фазовые искажения сравнительно не важны. Радиолюбители также используют системы с ОБП.

---

[1]) Применение квадратурного уплотнения (см. задачу 3.3) может устранить это различие. Однако этот способ вносит бо́льшие переходные помехи (междуканальную интерференцию) вследствие неидеальности канала ДБП, чем те, которые свойственны частотному уплотнению в случае ОБП.

Однако получение сигналов ОБП большой мощности значительно труднее, чем сигналов ДБП (см. § 3.4). Этот недостаток устраняется в системах с частично-подавленной боковой полосой. Действительно, такой способ передачи является компромиссом между ОБП и ДБП; он объединяет преимущества обеих систем и устраняет их недостатки.

3.8. Амплитудная модуляция с частично-подавленной боковой полосой

Как уже отмечалось (§ 3.3) сигналы ОБП сравнительно трудно получать. Например, для получения сигнала ОБП посредством отфильтровки одной из боковых полос сигнала ДБП требуется фильтр с резкой отсечкой, что практически сложно реализовать. Указанная трудность преодолевается частичным подавлением боковой полосы. Этот способ передачи предусматривает плавную отсечку одной из боковых полос (рис. 3.22*г*), а не резкую, как при ОБП (рис. 3.22*в*). Частотная характеристика используемого в этом случае фильтра такова, что частичное подавление передаваемой боковой полосы (верхней боковой полосы на рис. 3.22*г*) вблизи несущей частоты в точности

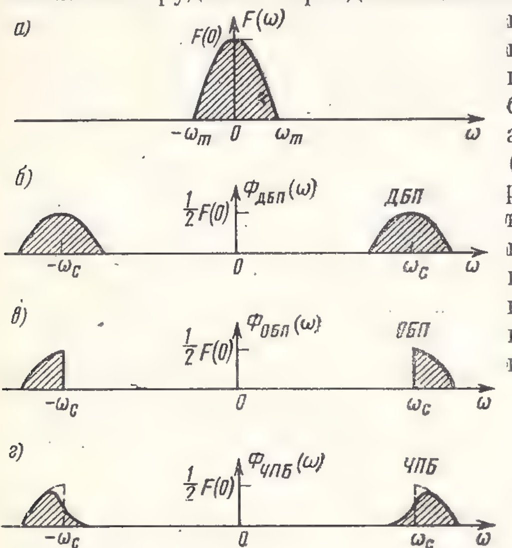


Рис. 3.22. Сравнение спектров сигналов при различных видах амплитудной модуляции: *а*) спектр модулирующего сообщения $f(t)$; *б*) спектр сигнала ДБП-ПН; *в*) спектр сигнала ОБП-ПН; *г*) спектр сигнала с частично-подавленной боковой полосой (ЧПБ)

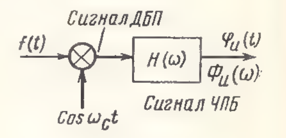


Рис. 3.23. Структурная схема получения сигналов с частично-подавленной боковой полосой

компенсируется частичной передачей подавляемой боковой полосы (нижней боковой полосы на рис. 3.22*г*) и полезное сообщение можно без искажений выделить соответствующим детектором.

Если вместе с сигналом с частично-подавленной боковой полосой передается несущее колебание большой амплитуды, детектирование может осуществляться детектором огибающей (или детектором-выпрямителем). Если же несущая отсутствует, сигнал может детектироваться синхронным детектором (или путем восста-

новления несущего колебания с последующим применением детектора огибающей).

Определим характеристику фильтра, которая требуется для получения сигнала с частично-подавленной боковой полосой из сигнала ДБП. Пусть $H(\omega)$ — передаточная функция искомого фильтра (рис. 3.23). Если $f(t)$ — сообщение, а $F(\omega)$ — его преобразование Фурье, то преобразование Фурье $\Phi_v(\omega)$ сигнала с частично-подавленной боковой полосой $\varphi_v(t)$:

$$\Phi_v(\omega)=\frac{1}{2}\,[F(\omega+\omega_c)+F(\omega-\omega_c)]\,H(\omega). \tag{3.45}$$

Сообщение $f(t)$ можно выделить из сигнала $\varphi_v(t)$ посредством синхронного детектирования. Для этого сначала входной сигнал $\varphi_v(t)$ умножается на $\cos\omega_c t$. Произведение $u_d(t)$ имеет преобразование Фурье

$$u_d(t)=\varphi_v(t)\cos\omega_c t \longleftrightarrow \frac{1}{2}\,[\Phi_v(\omega+\omega_c)+\Phi_v(\omega-\omega_c)].$$

Подставив в это выражение ф-лу (3.45), получим

$$u_d(t)\longleftrightarrow \frac{1}{4}\,\{[F(\omega+2\omega_c)+F(\omega)]H(\omega+\omega_c)+[F(\omega)+F(\omega-2\omega_c)]H(\omega-\omega_c)\}.$$

Слагаемые $F(\omega+2\,\omega_c)$ и $F(\omega-2\,\omega_c)$ представляют собой спектр $F(\omega)$, сдвинутый на частоты $\pm 2\,\omega_c$; они отфильтровываются фильтром нижних частот. Напряжение на выходе фильтра $u_o(t)$ имеет преобразование Фурье

$$u_o(t)\longleftrightarrow \frac{1}{4}\,F(\omega)\,[H(\omega+\omega_c)+H(\omega-\omega_c)]. \tag{3.46}$$

Для неискаженного детектирования необходимо, чтобы

$$u_o(t)\longleftrightarrow kF(\omega). \tag{3.47}$$

Следовательно, передаточная функция $H(\omega)$ должна удовлетворять условию

$$H(\omega+\omega_c)+H(\omega-\omega_c)=C. \tag{3.48}$$

Поскольку $F(\omega)=0$ при $|\omega|>\omega_m$, из ф-лы (3.46) следует, что условие (3.48) должно выполняться только для $|\omega|<\omega_m$. Слагаемые $H(\omega+\omega_c)$ и $H(\omega-\omega_c)$ представляют собой передаточную функцию $H(\omega)$, сдвинутую относительно начала координат на $-\omega_c$ и $+\omega_c$ соответственно (рис. 3.24*б* и *в*). Сумма этих двух слагаемых должна быть постоянна в интервале частот $|\omega|<\omega_m$. Это возможно только в том случае, если характеристика фильтра относительно несущей частоты обладает симметрией, показанной на рис. 3.24*д*.

Можно показать, что если к сигналу с частично-подавленной боковой полосой добавить несущую с большой амплитудой, то детектирование можно выполнить и посредством детектора огибаю-

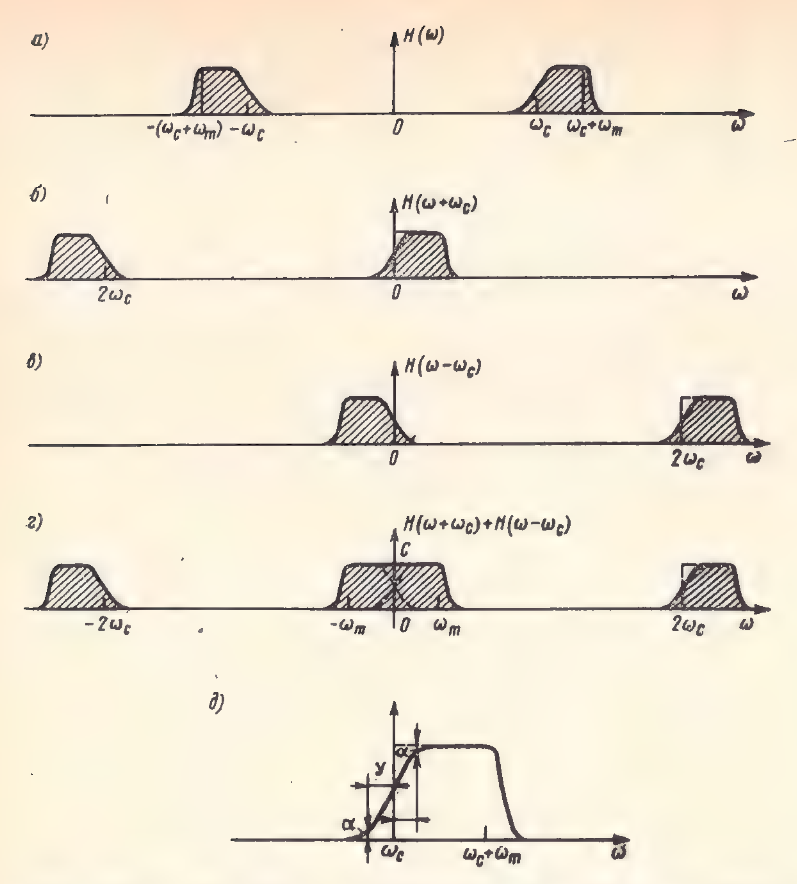


Рис. 3.24. К определению передаточной функции фильтра, частично подавляющего боковую полосу частот:
а) передаточная функция $H(\omega)$; *б)* слагаемое $H(\omega+\omega_c)$; *в)* слагаемое $H(\omega-\omega_c)$; *г)* передаточная функция детектора сигнала ЧПБ: *д)* передаточная функция $H(\omega)$, обладающая центральной симметрией

щей (или детектора-выпрямителя). Если амплитуда несущей велика по сравнению с $\varphi_v(t)$, то суммарный сигнал будет иметь пересечения нулевого уровня почти в тех же точках, что и несущее колебание и, следовательно, выпрямление такого сигнала будет эквивалентно умножению его на прямоугольное периодическое колебание $p(t)$. Очевидно, это эквивалентно синхронному детектированию.

Система с частично-подавленной боковой полосой обладает преимуществами систем ОБП и ДБП, не имея свойственных им недостатков. Эта система занимает практически ту же полосу частот, что и ОБП (половину полосы ДБП) и может получаться из сигна-

ла ДБП с помощью относительно простых фильтров с плавной характеристикой затухания. Она сравнительно нечувствительна к селективным замираниям. Если в сигнал с частично-подавленной боковой полосой ввести несущее колебание большой амплитуды, такой сигнал можно детектировать детектором огибающей. В этом случае система с частично-подавленной боковой полосой объединяет все преимущества АМ, ОБП и ЧПБ. Она используется в вещательных телевизионных системах для передачи видеосигналов.

3.9. Системы связи с частотным уплотнением

Как уже указывалось, модуляция необходима для одновременной передачи нескольких сообщений. Пусть требуется по одному каналу одновременно передавать n сообщений, каждое из которых имеет спектр, ограниченный частотой ω_m, *рад/сек*. В качестве примера рассмотрим использование АМ. Эти n сообщений модулируют несущие колебания с частотами $\omega_1, \omega_2, \ldots, \omega_n$, причем каждая несущая отличается от соседней, по крайней мере, на $2\omega_m$, *рад/сек*. Каждый из модулированных сигналов имеет ширину спектра $2\omega_m$ на частотах вблизи $\omega_1, \omega_2, \ldots, \omega_n$ (такие же спектры имеются на отрицательных частотах). Спектры сообщений показаны на рис. 3.25*а*, а на рис. 3.25*б* показан общий спектр всех модулирован-

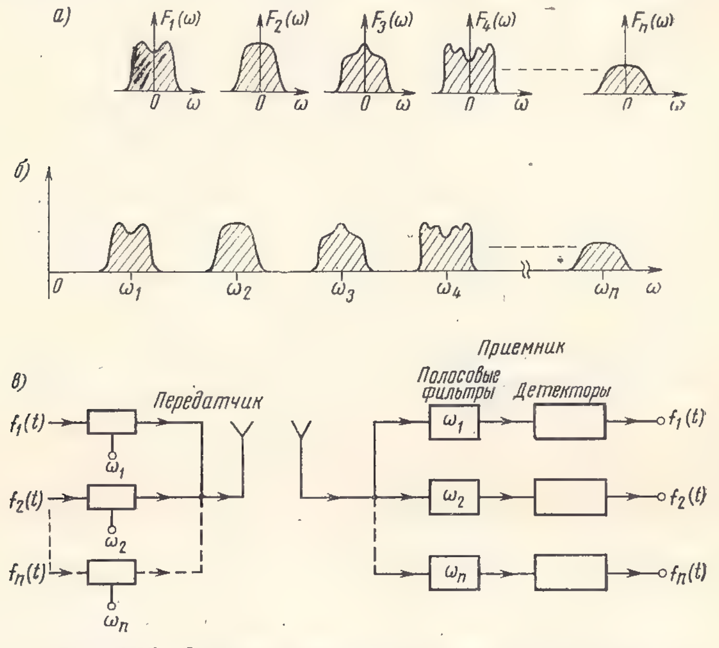


Рис. 3.25. Система связи с частотным уплотнением: *а*) спектры передаваемых сообщений $f_k(t)$; *б*) модулированные сигналы с неперекрывающимися спектрами; *в*) передатчик и приемник

ных сигналов в передатчике. На приемной стороне спектры различных сигналов выделяются соответствующими полосовыми фильтрами (рис. 3.25*в*). После фильтрации сигналы детектируются для получения переданных сообщений.

Примерами систем с частотным разделением являются системы радиовещания и телевидения. Каждый передатчик излучает модулированный сигнал на несущей частоте, которая отличается от несущих частот других станций, по крайней мере, на $2\omega_m$. В системах радиовещания этот интервал составляет около 10 *кгц*. Радиовещательный приемник может принять любой из сигналов при соответствующей настройке. Выделенный таким образом сигнал детектируется для получения передаваемого сообщения. Однако почти во всех радиовещательных АМ приемниках детектированию предшествует перенос спектра сигнала на фиксированную более низкую частоту, которая называется промежуточной частотой и равна 455 *кгц* [1]).

Полученный сигнал промежуточной частоты усиливается и детектируется для получения передаваемого сообщения.

Преимущество способа приема с преобразованием частоты заключается в том, что для приема различных станций необходимо перестраивать только первый каскад (и местный генератор) приемника. Необходимое усиление сигнала достигается на фиксированной промежуточной частоте, и усилитель промежуточной частоты не нуждается в перестройке. Процесс преобразования частоты называется также *гетеродинированием*. Чтобы переместить спектр сигнала на фиксированную промежуточную частоту, местный генератор должен иметь частоту, отличающуюся от несущей входного сигнала (в ту или иную сторону) на значение промежуточной частоты (455 *кгц*). Обычно частота местного генератора выбирается выше частоты входного сигнала. Поэтому такие приемники называются *супергетеродинными*. Структурная схема такого приемника показана на рис. 3.26.

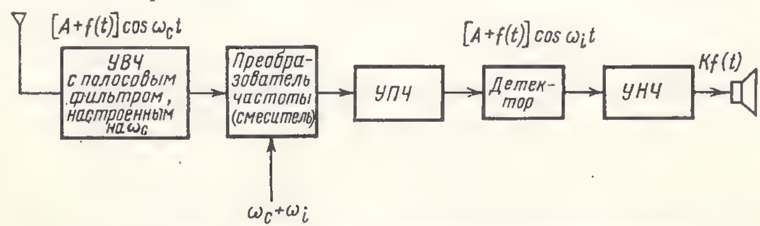


Рис. 3.26. Структурная схема супергетеродинного приемника

ЗАДАЧИ

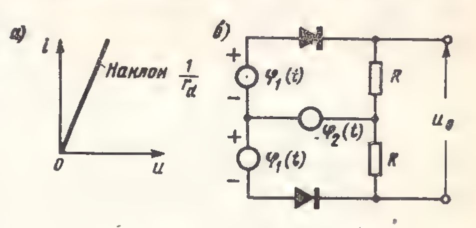
3.1. Покажите, что балансные модуляторы, схемы которых приведены на рис. 3.4, 3.5, 3.6, можно использовать для переноса спектра сигнала ДБП-ПН $\varphi(t)=f(t)\cos\omega_1 t$ на частоту ω_2, подавая колебания $f(t)\cos\omega_1 t$ и $A\cos\omega_2 t$ в со-

[1]) В отечественных радиовещательных приемниках промежуточная частота равна 465 *кгц*. — *Прим. ред.*

ответствующие точки схемы и применяя соответствующие фильтры. Предполагается, что $A \gg |f(t)|$.

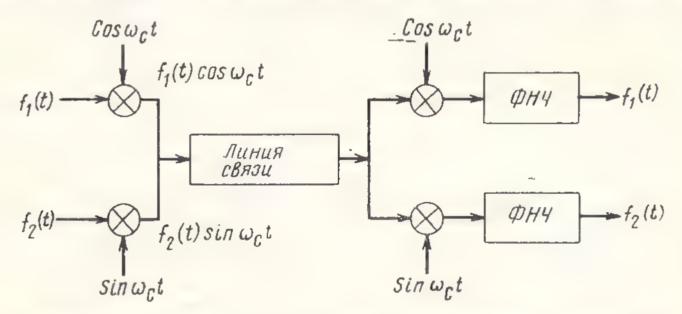
3.2. Балансный модулятор можно также использовать как синхронный детектор. Предполагая, что диоды имеют линейно-ломаную характеристику (см. рисунок), определите выходное напряжение $u_0(t)$ в схеме, показанной на рисунке, если $\varphi_1(t)=f(t)\cos\omega_c t$, $\varphi_2(t)=A\cos\omega_c t$. Предполагается, что $A \gg |f(t)|$ во все моменты времени.

Как можно выделить $f(t)$ из сигнала $u_0(t)$?

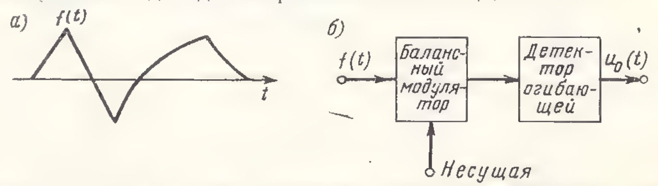


Эта же схема используется в качестве фазового дискриминатора, который измеряет фазовый сдвиг одного гармонического колебания относительно другого. Покажите, что если $\varphi_1(t)=\cos(\omega t+\theta)$, $\varphi_2(t)=A\cos\omega t$, то выходное напряжение $u_0(t)$ содержит постоянную составляющую, пропорциональную $\cos\theta$.

3.3. Возможна одновременная передача двух различных сообщений на одной несущей частоте. Два сообщения модулируют несущие с одинаковыми частотами, но сдвинутые по фазе на $\pi/2$, как показано на рисунке. Покажите, что можно выделить передаваемые сообщения с помощью синхронного детектирования принимаемого сигнала [1]).



3.4. Объясните, что получится, если детектор огибающей (или детектор-выпрямитель) использовать для детектирования сигналов ДБП-ПН.



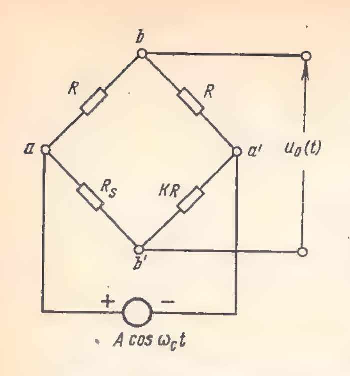
Сообщение $f(t)$ (см. рисунок) передается сигналом ДБП-ПН. Модулированный сигнал поступает на детектор огибающей. Представьте форму выходного напряжения детектора.

3.5. Модулирующее сообщение определяется выражением $f(t)=A\cos(2000\pi t)$. Представьте графически сигналы ДБП-ПН и ОБП-ПН, если несущая частота равна 10 *кгц*. Изобразите сигнал АМ, если коэффициент модуляции $m=0{,}75$.

3.6. В практике измерений часто используется мостовая цепь. Мост из резисторов применяется для измерения физических величин, действие которых приводит к пропорциональному изменению сопротивления. Примером может служить датчик давления, сопротивление которого изменяется пропорционально давлению.

---

[1]) Этот способ передачи называется *квадратурным уплотнением*.



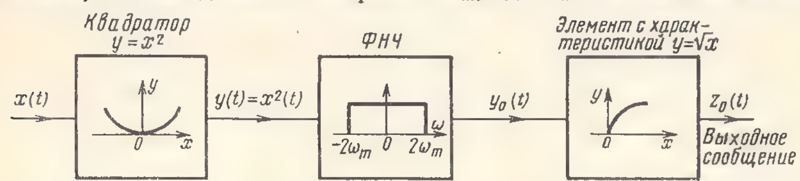
Подобным же образом сопротивление некоторых элементов (например, термисторов) изменяется пропорционально температуре. Мост, показанный на рисунке, имеет три фиксированных резистора. Четвертый резистор — переменный, его сопротивление изменяется пропорционально измеряемой физической величине:

$$R_s = KR\,[1 + \alpha f(t)],$$

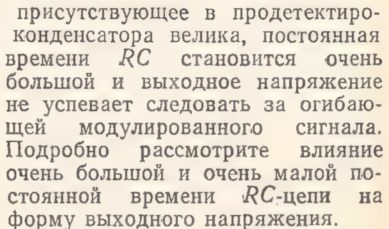
где KR — сопротивление в отсутствие воздействия; $f(t)$ — измеряемая величина (например, давление или температура); α — коэффициент пропорциональности.

К зажимам aa' подается напряжение $A\cos\omega_c t$. Определите выходной сигнал на зажимах bb' и укажите вид его модуляции. Как можно было бы детектировать такой сигнал? Предполагается, что $\alpha f(t) \ll 1$.

3.7. Покажите, что система, представленная на рисунке, может детектировать сигналы АМ. Определите частоту среза фильтра нижних частот *(ФНЧ)* и докажите, что она должна быть равна $2\omega_m$, где ω_m — наивысшая частота в

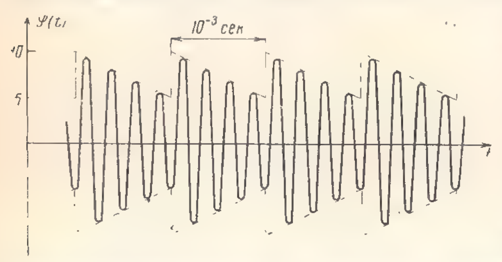


спектре сообщения $f(t)$. Покажите, что эта система действует как детектор огибающей и, следовательно, не может детектировать сигналы с подавленной несущей.



3.8. Покажите, что если в системе, рассмотренной в предыдущей задаче, первый каскад заменить линейным двухполупериодным выпрямителем с характеристикой, показанной на рисунке, то сигналы АМ можно детектировать, не включая в схему оконечный нелинейный элемент.

3.9. Конденсатор C в детекторе огибающей (рис. 3.14) должен быть достаточно большим, чтобы отфильтровывать напряжение несущей, присутствующее в продетектированном сигнале. Если, однако, емкость этого конденсатора велика, постоянная времени RC становится очень большой и выходное напряжение не успевает следовать за огибающей модулированного сигнала. Подробно рассмотрите влияние очень большой и очень малой постоянной времени RC-цепи на форму выходного напряжения.



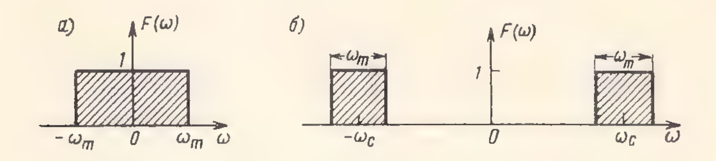
Определите наибольшее допустимое значение постоянной времени RC, чтобы выходное напряжение детектора следовало за огибающей модулированного сигнала, показанного на рисунке. Предполагается, что период модулирующего сигнала равен 10^{-3} *сек,* а период несущего колебания во много раз меньше 10^{-3} *сек.*

Указание. Аппроксимируйте экспоненциальный закон разряда RC-цепи первыми двумя членами ряда Тейлора, а скорость разряда RC-цепи приравняйте скорости изменения огибающей.

Если бы модулирующее сообщение являлось синусоидой частоты ω_s, то чему было бы равно наибольшее значение постоянной времени RC-цепи, при котором выходное напряжение следует за огибающей?

3.10. Изобразите форму колебаний в точках aa', bb' и cc' модулятора-прерывателя, показанного на рис. 3.10*в*. Какие условия для амплитуды k синусоидального напряжения и модулирующего сообщения $f(t)$ должны выполняться, чтобы выходное напряжение в точках cc' было сигналом АМ? Приведите необходимые объяснения, предполагая, что полосовой фильтр имеет бесконечное входное сопротивление.

3.11. Пусть спектр сообщения $f(t)$ ограничен частотой ω_m, *рад/сек*. Ограничен ли спектр колебания $f^2(t)$? Если ограничен, то найдите наивысшую частоту его спектра. Что в общем случае можно сказать о спектре колебания $f^n(t)$?



Указание. Используйте теорему частотной свертки.

Спектры двух сигналов показаны на рисунке. Постройте графики спектров $f^2(t)$.

3.12. На рисунке показан периодический модулирующий сигнал $f(t)$. Какое минимальное значение должна иметь амплитуда несущего колебания при АМ? Для этого значения амплитуды несущего колебания постройте график сигнала АМ, если несущая частота равна 100 *кгц*. Определите спектр сигнала АМ. Постройте график сигнала ДБН-ПН и его спектр.

3.13. Синхронное детектирование сигнала АМ-ПН возможно посредством умножения принимаемого модулированного сигнала на периодическую последовательность прямоугольных импульсов $p(t)$ частоты ω_c, как показано на рисунке 3.7. В результате перемножения получается сигнал $p(t)f(t)\cos\omega_c t$. Найдите аналитическое выражение для функции спектральной плотности этого сигнала и покажите, что из него можно выделить сообщение $f(t)$, используя фильтр нижних частот.

Указание. Используйте ф-лы (3.8) и (3.11).

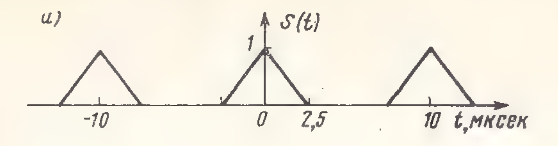
3.14. Некоторая станция использует сигнал ДБП-ПН средней мощности P, *вт*, чтобы обслужить определенную площадь. Какая мощность сигнала потребовалась бы для обслуживания такой же площади, если бы вместо сигнала ДБП-ПН использовался сигнал ОБП-ПН при одинаковой напряженности поля сигналов? Предполагается, что в обоих случаях используется синхронное детектирование при одинаковой амплитуде местной несущей.

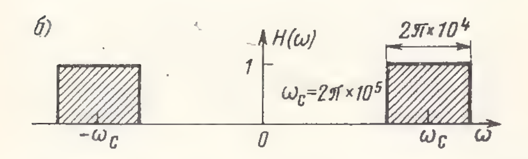
3.15. На рис. 3.12 показан однополупериодный диодный детектор-выпрямитель. Представьте схему двухполупериодного детектора такого типа и графики входного сигнала АМ, выпрямленного сигнала и напряжения на выходе фильтра нижних частот. Изобразите графики спектров этих колебаний.

Указание. Двухполупериодное выпрямление сигнала АМ эквивалентно умножению его на прямоугольное колебание с нулевым средним значением (биполярное колебание прямоугольной формы).

3.16. Искаженный гармонический сигнал $\cos^3\omega_c t$ используется как несущее колебание для получения сигнала ДБП-ПН. Найдите спектр произведения $f(t)\cos^3\omega_c t$ и постройте его график. Как из этого колебания выделить сигнал $f(t)\cos\omega_c t$?

3.17. Модулирующее сообщение $f(t)$, спектр которого ограничен частотой 5 *кгц*, умножается на периодическую последовательность треугольных импульсов (см. рисунок). Постройте график спектра произведения. Для получения мо-





дулированного сигнала произведение пропускается через идеальный полосовой фильтр с центральной частотой 100 *кгц*. Выходное напряжение этого фильтра равно $Af(t)\cos\omega_c t$. Определите постоянную A.

3.18. В сигнале АМ $\varphi_{АМ}(t) = (1+A\cos\omega_m t)\cos\omega_c t$, частота $\omega_m = 10\pi\times\times 10^3$ *рад/сек*, несущая частота $\omega_c=2\pi\cdot 10^5$ *рад/сек*, $A=15$. Можно ли детектировать этот сигнал детектором огибающей? Какое колебание получится на выходе детектора огибающей? Определите его спектр.

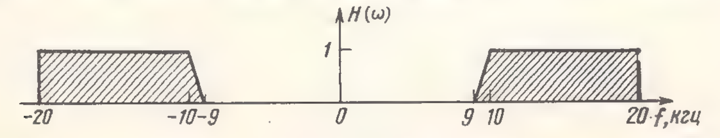
3.19. Покажите, что выходное напряжение в схеме рис. 3.10*б* действительно является сигналом АМ, если составной элемент, образованный диодом и резистором R, имеет вольтамперную характеристику вида: $i=au+bu^2$, где i — ток диода, u — напряжение на элементе диод-резистор. Предполагается, что полосовой фильтр имеет бесконечное сопротивление.

3.20. Модулирующее сообщение $f(t)=\cos(2000\pi t)+\cos(4000\pi t)$. Запишите выражение для сигнала ОБП-ПН, если несущая частота равна 10 *кгц*.

3.21. Определите выходное напряжение фазовращателя, показанного на рис. 3.19, если на его входе действует сигнал $A\cos(\omega_c t+\theta)$.

3.22. Запишите выражение сигнала ОБП-ПН для периодического модулирующего сообщения $f(t)$ (см. рисунок к задаче 3.12), если несущая частота равна 100 *кгц*.

3.23. Сигнал с частично-подавленной боковой полосой получается, если пропустить сигнал АМ через фильтр, амплитудно-частотная характеристика которого показана на рисунке. Найдите выражение для сигнала с частично-подав-



ленной боковой полосой, если модулирующее сообщение а) $f(t)=A\sin(100\pi t)$, б) $f(t)=A[\sin(100\pi t)+\cos(200\pi t)]$, в) $f(t)=A\sin(100\pi t)\cos(200\pi t)$.

Несущая частота равна 10 *кгц*, амплитуда несущей — $4A$. Постройте график спектра, получающего в каждом случае сигнала с частично-подавленной боковой.

3.24. Постройте график сигнала с частично-подавленной боковой полосой, если модулирующее сообщение в задаче 3.23 а) $f(t)=Sa(200\pi t)$, б) $f(t)=Sa^2(200\pi t)$.

Предполагается, что несущая подавлена.

ГЛАВА

4

Системы связи с угловой модуляцией

4.1. Угловая модуляция

В сигналах АМ амплитуда несущего колебания модулируется сообщением $f(t)$ и, следовательно, информация, содержащаяся в $f(t)$, переносится изменениями амплитуды несущей. Поскольку синусоидальное колебание зависит от трех переменных — амплитуды, частоты и фазы, — существует возможность передачи той же информации изменением либо частоты, либо фазы несущей. Синусоидальное колебание, однако, по определению представляет собой бесконечное колебание с постоянными амплитудой, частотой и фазой.

Поэтому необходимо распространить понятие синусоидального колебания на функцию общего вида, амплитуда, частота и фаза которой могут изменяться во времени. Понятие переменной амплитуды было введено ранее при изучении сигналов АМ. Теперь рассмотрим изменения частоты и фазы.

Чтобы лучше понять смысл изменений частоты, необходимо определить мгновенную частоту. На рис. 4.1*а* показан синусоидальный сигнал $\varphi(t)$, который имеет постоянную частоту ω_0 при $t<T$. В момент $t=T$ частота скачком изменяется до значения $2\omega_0$ и остается такой до $t=2T$, а затем снова изменяется до значения ω_0. Таким образом, функция $\varphi(t)$ есть синусоидальный сигнал, который имеет частоту ω_0 в интервалах времени $2nT<t<(2n+1)T$ и $2\omega_0$ в интервалах времени $(2n+1)T<t<(2n+2)T$ (n — целое число).

Рассмотрим теперь плавное изменение частоты, как показано на рис. 4.1*г*. Здесь частота сигнала изменяется непрерывно с постоянной скоростью от значения ω_0 до $2\omega_0$ в интервале T. Следовательно, в каждый момент времени частота разная. Строго говоря, сигнал $\varphi(t)$ на рис. 4.1*в* нельзя выразить через исходное синусоидальное колебание. Как же быть в случае непрерывных изменений частоты синусоидального сигнала? Для этого необходимо ввести определение обобщенного синусоидального сигнала

$$f(t)=A\cos\theta(t), \tag{4.1}$$

где θ — фаза синусоидального сигнала, являющаяся функцией времени.

В случае сигнала с постоянной частотой

$$f(t) = A\cos(\omega_c t + \theta_0).$$

Следовательно, $\theta(t) = \omega_c t + \theta_0$ и

$$\omega_c = d\theta/dt. \tag{4.2}$$

Здесь частота ω_c постоянна и определяется как производная угла $\theta(t)$. В общем случае эта производная может не быть постоянной. Назовем $d\theta/dt$ мгновенной частотой ω_i, которая может изменяться во времени. Таким образом, между углом $\theta(t)$ и мгновенной частотой имеет место соотношение

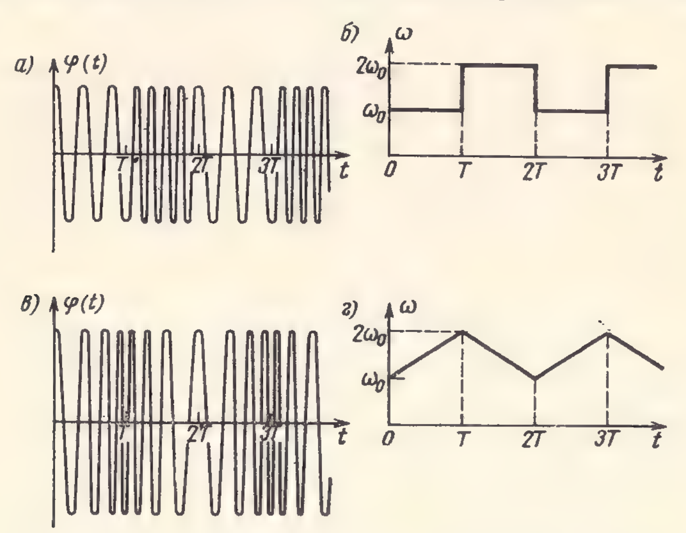


Рис. 4.1. К определению мгновенной частоты сигнала: *а)* сигнал со скачкообразным изменением частоты; *б)* скачкообразное изменение частоты; *в)* сигнал с непрерывным изменением частоты; *г)* непрерывное изменение частоты

$$\omega_i = d\theta/dt, \quad \theta = \int \omega_i\, dt. \tag{4.3}$$

Теперь легко представить, как можно передавать информацию, содержащуюся в сообщении $f(t)$, посредством изменения фазы несущей θ. Модуляция, при которой фаза несущей изменяется некоторым образом как функция $f(t)$, называется *угловой*. Обычно используются две разновидности угловой модуляции: *фазовая* (ФМ) и *частотная* (ЧМ). Если угол $\theta(t)$ изменяется прямо пропорционально $f(t)$, то

$$\theta(t) = \omega_c t + \theta_0 + k_p f(t). \tag{4.4}$$

Здесь k_p — некоторая постоянная. Такое изменение угла называется фазовой модуляцией. Таким образом, сигнал $A\cos[\omega_c t + \theta_0 +$

$+k_p f(t)$ является несущей, модулированной по фазе. Заметим, что мгновенная частота ω_i сигнала ФМ определяется выражением

$$\omega_i = d\theta/dt = \omega_c + k_p df/dt. \tag{4.5}$$

Следовательно, при фазовой модуляции мгновенная частота изменяется прямо пропорционально производной модулирующей функции. Если же пропорционально модулирующей функции $f(t)$ изменяется мгновенная частота, то получается частотная модуляция. Таким образом, для модулированной по частоте несущей

$$\omega_i = \omega_c + k_f f(t). \tag{4.6}$$

$$\theta(t) = \int \omega_i dt = \omega_c t + k_f \int f(t)\, dt + \theta_0. \tag{4.7}$$

Модулированная по частоте несущая записывается как $A\cos[\omega_c t + \theta_0 + k_f \int f(t) dt]$.

Из изложенного следует, что хотя ФМ и ЧМ являются разными формами угловой модуляции, различие между ними не столь существенно. При ФМ фаза изменяется линейно с модулирующей функцией, тогда как при ЧМ фаза изменяется линейно с интегралом модулирующей функции. Если сначала проинтегрировать модулирующую функцию $f(t)$, а затем этим колебанием модулировать несущую по фазе, то получится сигнал ЧМ. Подобным же образом, если продифференцировать $f(t)$ и это колебание использовать для модуляции частоты, то получится сигнал ФМ. Действительно, при получении сигналов ЧМ одним из возможных способов сообщение $f(t)$ интегрируется, а затем модулирует несущую по фазе. Фактически ЧМ и ФМ неотделимы, поскольку любые изменения фазы несущей приводят к изменениям частоты и наоборот. Поэтому здесь будет рассмотрена только частотная модуляция, однако все полученные выводы в равной степени справедливы и для фазовой модуляции. Заметим, что при угловой модуляции амплитуда сигнала остается постоянной. Примеры сигналов ЧМ показаны на рис. 4.1*а* и *в*.

Если сигналы ФМ и ЧМ обозначены соответственно $\varphi_{ФМ}(t)$ и $\varphi_{ЧМ}(t)$, то

$$\left.\begin{aligned} \varphi_{ФМ}(t) &= A\cos[\omega_c t + k_p f(t)] \\ \varphi_{ЧМ}(t) &= A\cos[\omega_c t + k_f \int f(t)\, dt] \end{aligned}\right\}, \tag{4.8}$$

где $f(t)$ — модулирующее сообщение. Начальная фаза θ_0 в этих формулах полагается равной нулю без потери общности.

Преимущества комплексного представления гармонических сигналов хорошо известны. Поэтому запишем гармонический сигнал (4.1) в комплексной форме

$$A\cos\theta(t) \sim A\,e^{i\theta(t)},$$

причем реальный сигнал есть действительная часть комплексного

$$A\cos\theta(t) = \mathrm{Re}\left[A\,e^{i\theta(t)}\right].$$

Для комплексного представления колебания $\varphi(t)$ будем использовать обозначение $\hat{\varphi}(t)$. Таким образом, если $\varphi(t) = A\cos\theta(t)$, то $\hat{\varphi}(t) = Ae^{i\theta(t)}$ и $\varphi(t) = \mathrm{Re}\,\hat{\varphi}(t)$.

В принятом обозначении сигналы ЧМ и ФМ [ф-ла (4.8)] записываются как

$$\left.\begin{aligned} \hat{\varphi}_{ФМ}(t) &= A\,e^{i[\omega_c t + k_p f(t)]} \\ \hat{\varphi}_{ЧМ}(t) &= A\,e^{i[\omega_c t + k_f \int f(t)\,dt]} \end{aligned}\right\}. \tag{4.9}$$

Введем обозначение $\int f(t)dt = g(t)$, тогда

$$\hat{\varphi}_{ЧМ}(t) = A\,e^{i[\omega_c t + k_f g(t)]}. \tag{4.10}$$

4.2. Узкополосная частотная модуляция

Общее выражение сигнала ЧМ дается ф-лой (4.10). Мгновенная частота

$$\omega_i = d\,\theta/dt = \omega_c + k_f dg/dt = \omega_c + k_f f(t)$$

прямо пропорциональна сообщению $f(t)$. Слагаемое $k_f f(t)$ представляет собой отклонения частоты от ее значения ω_c, причем постоянная k_f определяет величину отклонения частоты. При малых значениях k_f изменения частоты невелики, и следует ожидать, что спектр сигнала ЧМ будет занимать узкую полосу частот. Если постоянная k_f велика, то следует ожидать, что и спектр сигнала ЧМ окажется широким. Это действительно так. Рассмотрим сначала узкополосную ЧМ. Если постоянная k_f столь мала, что для всех t выполняется условие $k_f g(t) \ll 1$, то

$$e^{i k_f g(t)} \approx 1 + i\,k_f g(t), \tag{4.11}$$

$$\hat{\varphi}_{ЧМ}(t) \approx A\,[1 + i\,k_f g(t)]\,e^{i\omega_c t}, \tag{4.12}$$

$$\varphi_{ЧМ}(t) = \mathrm{Re}\,[\hat{\varphi}_{ЧМ}(t)] = A\cos\omega_c t - Ak_f g(t)\sin\omega_c t. \tag{4.13}$$

В полученном выражении первое слагаемое представляет собой несущую, а второе — боковые полосы.

Интересно заметить, что сигнал АМ записывался в виде [ф-ла (3.12)]

$$\varphi_{АМ}(t) = A\cos\omega_c t + f(t)\cos\omega_c t,$$

тогда как сигнал узкополосной ЧМ выражается формулой

$$\varphi_{ЧМ}(t) = A\cos\omega_c t - A\,k_f g(t)\sin\omega_c t.$$

Подобным же образом сигнал узкополосной ФМ

$$\varphi_{ФМ}(t) = A\cos\omega_c t - Ak_p f(t)\sin\omega_c t. \tag{4.14}$$

Каждый из этих сигналов содержит несущее колебание и боковые полосы вблизи частот $\pm\omega_c$. Если $f(t)\longleftrightarrow F(\omega)$ и $g(t)\longleftrightarrow G(\omega)$, то поскольку $g(t)=\int f(t)dt$, из ф-лы (1.119б) следует, что

$$g(t)\longleftrightarrow \frac{1}{i\omega}F(\omega). \quad (4.15)$$

Следовательно,

$$G(\omega)=\frac{1}{i\omega}F(\omega). \quad (4.16)$$

Таким образом, если спектр $F(\omega)$ ограничен частотой ω_m, то спектр $G(\omega)$ также ограничен частотой ω_m. Спектр сигнала ЧМ (4.13) можно найти, используя ф-лу (1.116б). Итак, если $\varphi_{ЧМ}(t)\longleftrightarrow\Phi_{ЧМ}(\omega)$, то

$$\Phi_{ЧМ}(\omega)=\pi A[\delta(\omega-\omega_c)+\delta(\omega+\omega_c)]+\frac{iAk_f}{2}[G(\omega-\omega_c)-G(\omega+\omega_c)]. \quad (4.17)$$

Сравнение спектров сигнала ЧМ [ф-ла (4.17)] и сигнала АМ [ф-ла (3.13)] позволяет установить сходство и различие между этими двумя видами модуляции. Оба вида модуляции имеют несущие на частотах $\pm\omega_c$. Однако боковые спектры в случае ЧМ имеют фазовый сдвиг $\pi/2$ относительно несущего колебания, тогда как при АМ боковые находятся в фазе с несущей. Спектр $G(\omega)=(1/i\omega)F(\omega)$, и, следовательно, если спектр $F(\omega)$ ограничен частотой ω_m, то $G(\omega)$ также ограничен частотой ω_m. Таким образом,

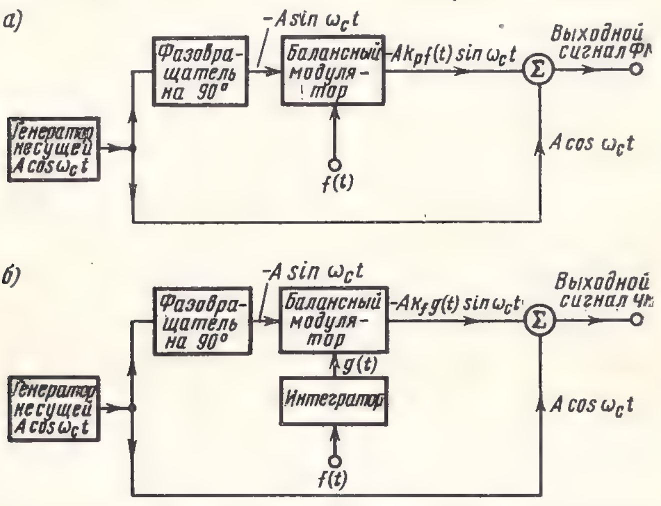


Рис. 4.2. Структурные схемы получения сигналов узкополосных ФМ и ЧМ с использованием балансных модуляторов: *а)* фазовый модулятор; *б)* частотный модулятор

сигналы узкополосной ЧМ (и ФМ) занимают ту же ширину спектра $2\omega_m$, что и сигнал АМ. Следует, однако, напомнить, что, несмотря на отмеченное сходство, сигналы АМ и ЧМ имеют совершенно различную форму. В сигнале АМ частота постоянна, а изменяется амплитуда, тогда как в сигнале ЧМ амплитуда постоянна, а частота изменяется.

Формулы (4.13) и (4.14) указывают возможный способ получения сигналов узкополосной ЧМ и ФМ при помощи балансных модуляторов. Структурные схемы частотного и фазового модуляторов показаны на рис. 4.2.

4.3. Широкополосная частотная модуляция

Если отклонения частоты при модуляции достаточно велики, т. е. если постоянная k_f выбрана столь большой, что условие $k_f g(t) \ll 1$ не выполняется, то анализ сигналов ЧМ в общем случае при произвольной функции $f(t)$ становится очень сложным. Получение точного выражения для ширины спектра сигнала ЧМ в случае произвольной модулирующей функции $f(t)$ невозможно, так как частотная модуляция является нелинейной (см. § 4.4 и 4.6). Поэтому сначала получим выражение для ширины спектра сигнала ЧМ на эвристической основе, а затем докажем этот результат для некоторых модулирующих функций. Покажем, что ширина спектра сигнала ЧМ приближенно выражается формулой

$$W = 2\left[k_f \left|f(t)\right|_{\text{макс}} + 2\omega_m\right],$$

где ω_m — ширина спектра сообщения $f(t)$.

Заметим, что $\omega_i = \omega_c + k_f f(t)$. Следовательно, $k_f \left|f(t)\right|_{\text{макс}}$ представляет наибольшее отклонение частоты. Если это произведение обозначить через $\Delta\omega$, то

$$W = 2\left[\Delta\omega + 2\omega_m\right]. \tag{4.18}$$

Чтобы получить этот результат, сообщение $f(t)$ приближенно представим в виде ступенчатой функции, как показано на рис. 4.3*а*. Сообщение $f(t)$ имеет спектр, ограниченный частотой f_m, *гц*. Поэтому в соответствии с теоремой отсчетов оно будет почти постоянно на интервале $1/(2f_m)$, *сек* (рис. 4.3*а*). Сигнал ЧМ, соответствующий ступенчатой функции, будет состоять из импульсов, частота которых постоянна на интервале $1/(2f_m)$, *сек*.

Заметим, что на границах интервалов частота изменяется скачком. Один из таких импульсов показан на рис. 4.3*б*. Спектр каждого импульса можно получить, используя ф-лу (6) из табл. 1.1 и теорему модуляции (см. также рис. 1.22). Из рис. 4.3*в* видно, что импульс занимает полосу частот от $\omega_i - 2\omega_m$ до $\omega_i + 2\omega_m$, где ω_i — частота заполнения импульса. В данном случае $\omega_i = \omega_c + k_f f(t_k)$, где t_k — k-й момент отсчета. Следовательно, спектр лежит в области частот от $\omega_c + k_f f(t_k) - 2\omega_m$ до $\omega_c + k_f f(t_k) + 2\omega_m$. Очевидно, спектр

всего сигнала ЧМ лежит в диапазоне частот от $\omega_c - k_f|f(t)|_{\text{макс}} - 2\omega_m$ до $\omega_c + k_f|f(t)|_{\text{макс}} + 2\omega_m$ и ширина его

$$W \approx 2k_f\,|f(t)|_{\text{макс}} + 4\omega_m = 2(\Delta\omega + 2\omega_m).$$

В случае широкополосной ЧМ $\Delta\omega \gg \omega_m$ и $W \approx 2\Delta\omega$. Следовательно, ширина спектра сигнала широкополосной ЧМ равна примерно удвоенной девиации частоты $\Delta\omega$. Этот результат очевиден,

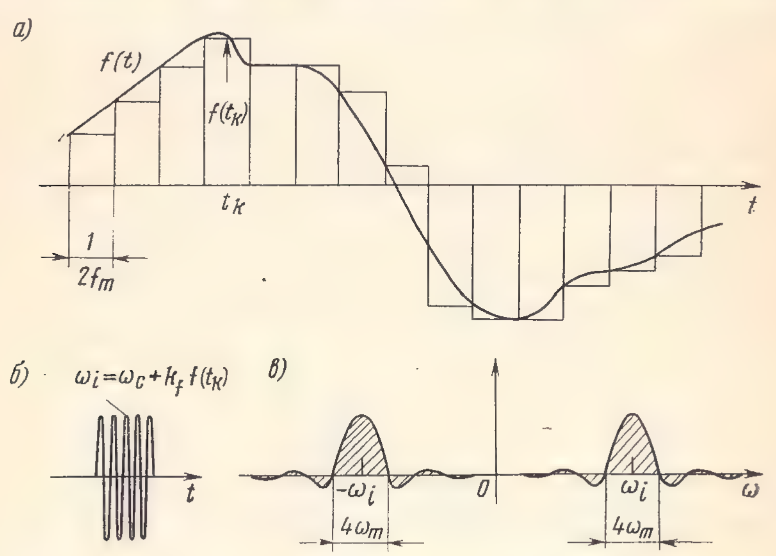


Рис. 4.3. К определению ширины спектра сигнала широкополосной ЧМ

так как если наибольшее отклонение частоты равно $\Delta\omega$, то частота сигнала ЧМ изменяется от значения $\omega_c - \Delta\omega$ до $\omega_c + \Delta\omega$. В первом приближении сигнал ЧМ содержит частоты, лежащие в этом диапазоне, а значит ширина его спектра равна $2\Delta\omega$.

Теперь подтвердим этот вывод для случая, когда модулирующее сообщение $f(t) = a\cos\omega_m t$.

Можно предположить, что сообщение $f(t)$ включается в момент $t=0$, и потому

$$g(t) = \int f(t)\,dt = a\int_0^t \cos\omega_m t\,dt = \frac{a}{\omega_m}\sin\omega_m t. \tag{4.19}$$

Поскольку мгновенная частота в соответствии с ф-лой (4.6)

$$\omega_i = \omega_c + k_f f(t) = \omega_c + ak_f\cos\omega_m t, \tag{4.20}$$

наибольшее отклонение частоты равно ak_f, *рад/сек:*

$$\Delta\omega = ak_f. \tag{4.21}$$

Подставляя ф-лу (4.19) в (4.10) и используя ф-лу (4.21), получаем

$$\hat{\varphi}_{ЧМ}(t) = A\,e^{i[\omega_c + (\Delta\omega/\omega_m)\sin\omega_m t]}.$$

Величина $\Delta\omega/\omega_m$, равная отношению наибольшего отклонения частоты $\Delta\omega$ к частоте модулирующей функции ω_m, называется *индексом модуляции* m_f. Таким образом,

$$ak_f/\omega_m = \Delta\omega/\omega_m = m_f \tag{4.22}$$

и

$$\hat{\varphi}_{ЧМ}(t) = A\,e^{i(\omega_c t + m_f \sin\omega_m t)} = A e^{i\,m_f \sin\omega_m t}\; e^{i\,\omega_c t}. \tag{4.23}$$

Первый экспоненциальный множитель в ф-ле (4.23) является, очевидно, периодической функцией с периодом $2\pi/\omega_m$ и может быть разложен в ряд Фурье

$$e^{i\,m_f \sin\omega_m t} = \sum_{n=-\infty}^{\infty} C_n e^{i\,n\omega_m t},$$

где

$$C_n = \frac{\omega_m}{2\pi} \int\limits_{-\pi/\omega_m}^{\pi/\omega_m} e^{i m_f \sin\omega_m t}\, e^{-i n\omega_m t}\, dt.$$

Обозначив $\omega_m t = x$, получим

$$C_n = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} e^{i(m_f \sin x - nx)}\, dx.$$

Этот интеграл в замкнутой форме не вычисляется, но может быть вычислен разложением подынтегральной функции в бесконечный ряд. Он обозначается как $J_n(m_f)$ и называется функцией Бесселя первого рода n-го порядка[1]). Графики этих функций приведены на рис. 4.4*а*. Таким образом,

$$e^{i\,m_f \sin\omega_m t} = \sum_{n=-\infty}^{\infty} J_n(m_f)\, e^{i\,n\omega_m t}. \tag{4.24}$$

Кроме того, можно показать, что

$$\left.\begin{aligned} J_n(m_f) &= J_{-n}(m_f), \text{ если } n \text{ четное} \\ J_n(m_f) &= -J_{-n}(m_f), \text{ если } n \text{ нечетное} \end{aligned}\right\}. \tag{4.25}$$

---

[1]) Таблицы функций Бесселя имеются, например, в книге Е. Янке, Ж. Эмде «Таблицы функций с формулами и кривыми» ГИТТЛ, 1949. — *Прим. ред.*

Подставив ф-лу (4.24) в (4.23), получим

$$\hat{\varphi}_{ЧМ}(t) = A\, e^{i\omega_c t} \sum_{n=-\infty}^{\infty} J_n(m_f)\, e^{in\omega_m t}$$

и

$$\varphi_{ЧМ}(t) = A \sum_{n=-\infty}^{\infty} J_n(m_f) \cos(\omega_c + n\,\omega_m)t. \qquad (4.26а)$$

Используя свойства (4.25), можно перписать ф-лу (4.26а) в виде

$$\varphi_{ЧМ}(t) = A\{J_0(m_f)\cos\omega_c t + J_1(m_f)[\cos(\omega_c+\omega_m)t - \cos(\omega_c-\omega_m)t] +$$
$$+ J_2(m_f)[\cos(\omega_c+2\omega_m)t + \cos(\omega_c-2\omega_m)t] + \ldots \qquad (4.26б)$$

Из ф-л (4.26) следует, что модулирующее сообщение $f(t)$, имеющее частоту ω_m, дает боковые частоты $\omega_c \pm \omega_m$, $\omega_c \pm 2\,\omega_m$, $\omega_c \pm 3\,\omega_m$ и т. д., как показано на рис. 4.4*б*. Сигнал ЧМ содержит составляющие с бесконечными частотами, и его спектр бесконечно широк. На практике, однако, амплитуды спектральных составляющих на высоких частотах весьма малы, и поэтому почти вся энергия сигнала ЧМ сосредоточена в составляющих, занимающих конечную полосу частот (рис. 4.4*а*). При $m_f \ll 1$ только $J_0(m_f)$ и $J_1(m_f)$ имеют существенное значение. Это случай узкополосной ЧМ, который рассматривался в § 4.1.

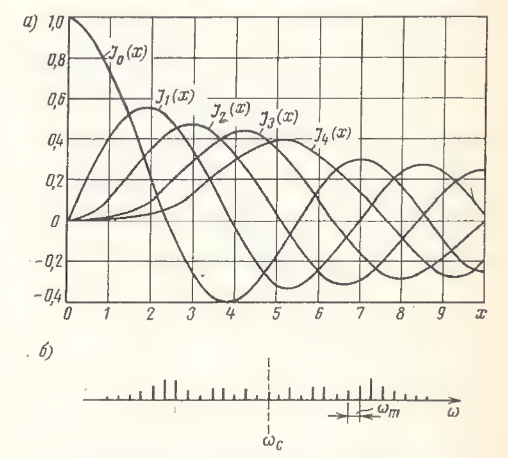


Рис. 4.4. К определению амплитуд спектральных составляющих сигнала ЧМ:
а) **бесселевы функции первого рода;**
б) **спектр сигнала ЧМ**

При $m_f=2$ функции $J_5(2)$, $J_6(2)$ и т. д. имеют пренебрежимо малые значения. Следовательно, значительными частотными составляющими при $m_f=2$ будут составляющие на частотах ω_c, $\omega_c \pm \omega_m$, $\omega_c \pm 2\,\omega_m$, $\omega_c \pm 3\,\omega_m$ и $\omega_c \pm 4\,\omega_m$. Они занимают полосу частот $8\,\omega_m$. При дальнейшем увеличении m_f возрастают амплитуды боковых частот более высокого порядка. Если учитываются лишь те составляющие, амплитуды которых превышают 1% от амплитуды немодулированной несущей, то для них выполняется условие $J_n(m_f) > 0{,}01$. Число существенных боковых составляющих при различных значениях m_f можно определить из графиков функций Бесселя. Из этих графиков видно, что $J_n(m_f)$ быстро убывает при

$n > m_f$. В общем случае $J_n(m_f)$ имеет пренебрежимо малые значения при $n > m_f$. Это приближение тем точнее, чем больше m_f. Таким образом, в случае широкополосной ЧМ, когда число существенных составляющих можно рассматривать равным целому числу, ближайшему к m_f, $m_f = n$. Общая ширина спектра сигнала ЧМ $W \approx \approx 2n\omega_m \approx 2m_f\omega_m$, m_f определяется по ф-ле (4.22).

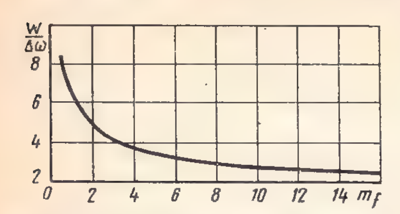


Рис. 4.5. Зависимость действительной ширины спектра сигнала ЧМ от индекса модуляции

Следовательно,

$$W \approx 2m_f\,\omega_m = 2\Delta\omega. \qquad (4.27)$$

Более точной оценкой является

$$W \approx 2\,(\Delta\omega + 2\omega_m). \qquad (4.28)$$

Для узкополосной ЧМ, при которой $\Delta\omega \ll \omega_m$, ширина спектра приближенно равна $2\,\omega_m$ (такая же, как для сигналов АМ). Следовательно, ширина полосы частот, которая требуется для передачи сигнала ЧМ, приближенно равна удвоенному наибольшему отклонению частоты. Точное соотношение между W и $\Delta\omega$ как функция m_f показано на рис. 4.5. Из этого рисунка видно, что при $m_f \gg 1$ ширина спектра W приближенно равна удвоенному значению девиации частоты.

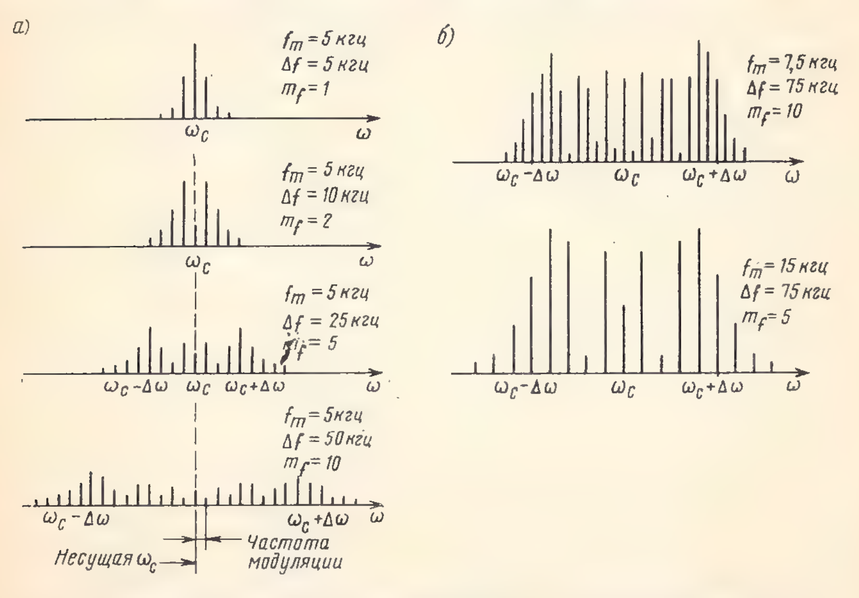


Рис. 4.6. Спектры сигналов ЧМ при различных значениях индекса и частоты модуляции:
а) частота f_m постоянна, индекс изменяется; *б)* девиация частоты постоянна, частота f_m изменяется

Для радиовещательных станций установлено наибольшее отклонение частоты, равное 75 *кгц*. Поэтому для передачи таких сигналов требуется полоса, приблизительно равная 150 *кгц*.

Из рис. 4.5 следует, что при широкополосной ЧМ ($m_f \gg 1$) требуемая для передачи полоса частот приближается к $2\Delta\omega$, *рад/сек* или $2\Delta f$, *гц*. Поскольку $\Delta\omega = ak_f = m_f\omega_m$, то при постоянном значении ω_m с ростом m_f увеличивается также и $\Delta\omega$. Это показано на рис. 4.6*а* для значений $m_f = 1$, 2, 5 и 10. Девиация частоты Δf равна соответственно 5, 10, 25 и 50 *кгц*. Заметим, что при больших значениях m_f ширина спектра сигнала равна приближенно $2\Delta f$.

На рис. 4.6*б* показан случай, когда девиация частоты сохраняется постоянной, а $m_f = \Delta\omega/\omega_m$ изменяется с изменением ω_m. Здесь $\Delta f = 75$ *кгц*, а m_f изменяется от 10 до 5 при изменении f_m от 7,5 до 15 *кгц*. В этом случае ширина спектра сигнала равна приближенно $B \approx 2\Delta f = 150$ *кгц*.

4.4. Частотная модуляция несколькими синусоидальными колебаниями

В § 4.3 рассматривался частный случай, когда модуляция частоты вызывалась одним синусоидальным сообщением. Распространим полученные выводы на случай сообщения, состоящего из нескольких частотных составляющих. Рассмотрим вначале сообщение, состоящее из двух составляющих:

$$f(t) = a_1 \cos \omega_1 t + a_2 \cos \omega_2 t.$$

Мгновенная частота

$$\omega_1 = \omega_c + k_f(t) = \omega_c + k_f(a_1 \cos \omega_1 t + a_2 \cos \omega_2 t).$$

Наибольшее отклонение частоты $\Delta\omega = (a_1 + a_2)k_f$ и

$$\theta(t) = \int \omega_i \, dt = \omega_c t + \frac{a_1 k_f}{\omega_1} \sin \omega_1 t + \frac{a_2 k_f}{\omega_2} \sin \omega_2 t = \omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t,$$

где

$$m_1 = a_1 k_f/\omega_1, \quad m_2 = a_2 k_f/\omega_2.$$

Комплексный сигнал ЧМ записывается в виде

$$\hat{\varphi}_{ЧМ}(t) = A e^{i\theta(t)} = A e^{i(\omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t)} =$$

$$= A e^{i\omega_c t} \left(e^{i m_1 \sin \omega_1 t}\right) \left(e^{i m_2 \sin \omega_2 t}\right). \qquad (4.29)$$

Заключенные в круглые скобки множители являются периодическими функциями с периодами $2\pi/\omega_1$ и $2\pi/\omega_2$ соответственно. Их можно разложить в ряд Фурье, используя функции Бесселя

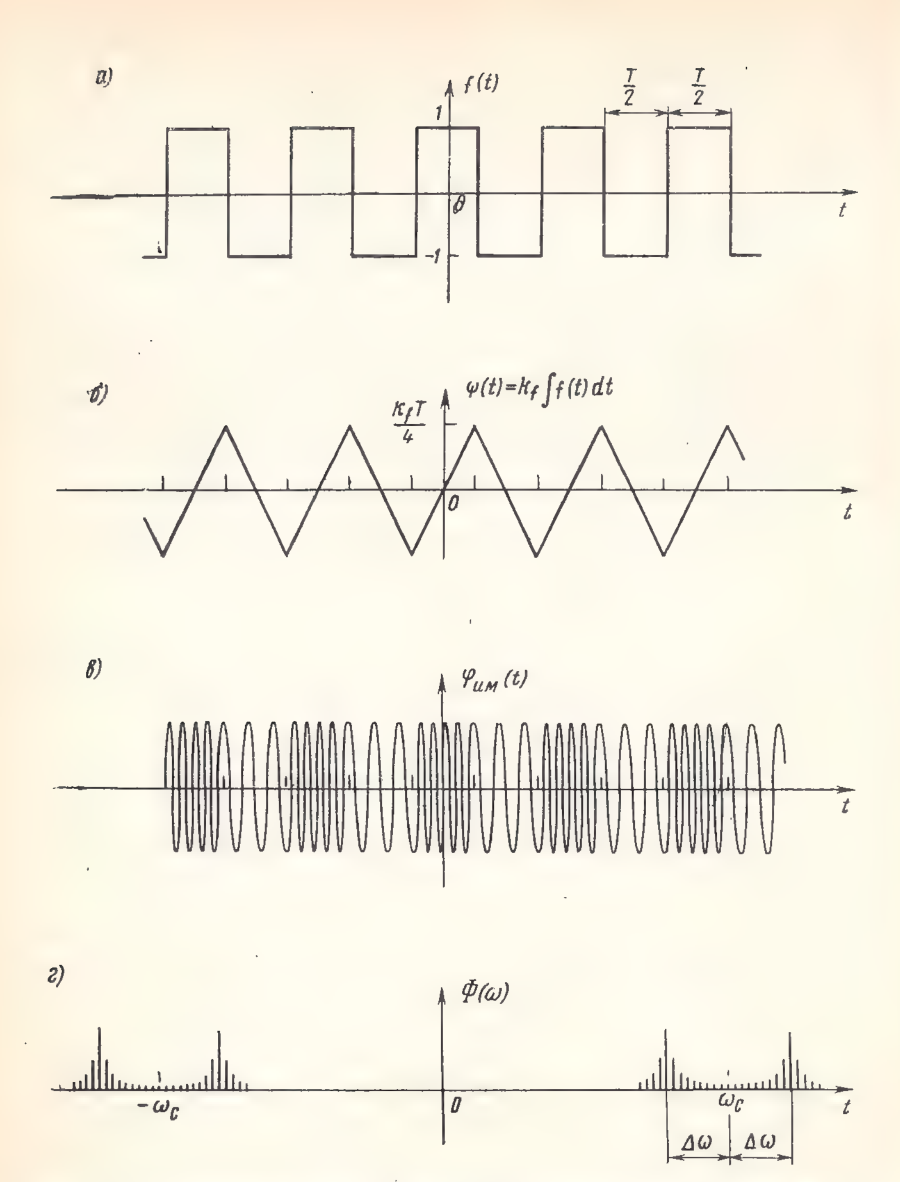


Рис. 4.7. К определению спектра сигнала ЧМ при модуляции прямоугольным колебанием:

а) модулирующее колебание; *б)* изменение фазы сигнала; *в)* сигнал частотной модуляции; *г)* спектр сигнала частотной модуляции

[ф-ла (4.24)]. Следовательно,

$$\hat{\varphi}_{ЧМ}(t) = A e^{i\omega_c t}\left[\sum_{n=-\infty}^{\infty} J_n(m_1) e^{i n\omega_1 t}\right]\left[\sum_{k=-\infty}^{\infty} J_k(m_2) e^{i k\omega_2 t}\right] =$$

$$= A \sum_{n=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} J_n(m_1) J_k(m_2) e^{i(\omega_c+n\omega_1+k\omega_2)t}. \quad (4.30)$$

и

$$\varphi_{ЧМ}(t) = A \sum_{n=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} J_n(m_1) J_k(m_2)\cos(\omega_c + n\,\omega_1 + k\,\omega_2)\,t. \quad (4.31)$$

Из этого выражения следует, что если $f(t)$ состоит из двух частот ω_1 и ω_2, спектр сигнала ЧМ содержит боковые полосы $(\omega_c \pm \pm n\omega_1)$ и $(\omega_c \pm k\omega_2)$. Кроме того, имеются составляющие комбинационных частот $(\omega_c \pm n\omega_1 \pm k\omega_2)$. Заметим, что ничего подобного не наблюдалось при АМ. В случае АМ каждая новая частота, добавляемая в модулирующую функцию, приводила к возникновению только своих собственных боковых частот. Комбинационные частоты отсутствовали. По этой причине амплитудная модуляция называется линейной, тогда как частотная — нелинейной. Более подробно этот вопрос рассматривается в § 4.6.

Полученные здесь выводы для двух частот легко обобщить на случай любого числа частот.

4.5. Модуляция колебанием прямоугольной формы

Остановимся на частном случае ЧМ, когда модулирующая функция $f(t)$ представляет собой прямоугольное колебание (рис. 4.7*а*). Рассматриваемый ниже метод является общим; он применим к периодическим модулирующим функциям любого вида. В случае прямоугольного колебания мгновенная частота и фаза сигнала ЧМ соответственно

$$\omega_i = \omega_c + k_f f(t), \quad \theta(t) = \int \omega_i\, dt = \omega_c t + k_f \int f(t)\, dt = \omega_c t + \psi(t).$$

График $\psi(t)$ показан на рис. 4.7*б*. Это треугольная периодическая функция с периодом T.

Заметим, что наибольшее отклонение частоты $\Delta\omega$ в этом случае равно k_f, так как $|f(t)|_{макс} = 1$. Следовательно, $k_f = \Delta\omega$ и

$$\psi(t) = \begin{cases} \Delta\omega\, t & \text{при} -\dfrac{T}{4} < t < \dfrac{T}{4}, \\ \Delta\omega\left(\dfrac{T}{2} - t\right) & \text{при} \quad \dfrac{T}{4} < t < \dfrac{3T}{4}, \end{cases} \quad (4.32)$$

а также $\psi(t) = \psi(t \pm nT)$.

В комплексной форме модулированный сигнал

$$\hat{\varphi}_{ЧМ}(t) = A e^{i\theta(t)} = A e^{i\psi(t)} e^{i\omega_c t}.$$

Функция $e^{i\psi(t)}$ — периодическая с периодом T и может быть разложена в ряд Фурье

$$e^{i\psi(t)} = \sum_{n=-\infty}^{\infty} a_n e^{in\omega_s t}; \quad \omega_s = 2\pi/T,$$

где $\alpha_n = \frac{1}{T} \int\limits_{-T/4}^{T/4} e^{i\psi(t)} e^{-in\omega_s t}\, dt.$

Подставив это выражение в ф-лу (4.32) и выполнив интегрирование, получим

$$\alpha_n = \frac{1}{2}\left\{Sa\left[\frac{\pi}{2}(\beta - n)\right] + (-1)^n Sa\left[\frac{\pi}{2}(\beta + n)\right]\right\}, \qquad (4.33)$$

где $\beta = \Delta\omega/\omega_s$.

Следовательно,

$$\hat{\varphi}_{ЧМ}(t) = A e^{i\psi(t)} e^{i\omega_c t} = A \sum_{n=-\infty}^{\infty} \alpha_n e^{i(\omega_c + n\omega_s)t},$$

$$\varphi_{ЧМ}(t) = A \sum_{n=-\infty}^{\infty} \alpha_n \cos(\omega_c + n\omega_s) t.$$

Спектр сигнала $\varphi_{ЧМ}(t)$ показан на рис. 4.7*г*.

Рассмотренный здесь метод является общим и применим к любым модулирующим функциям $f(t)$, которые являются периодическими и имеют нулевое среднее значение.

4.6. Линейная и нелинейная модуляции

ОБЩИЕ ПОЛОЖЕНИЯ

При амплитудной модуляции боковые частотные составляющие подчиняются принципу суперпозиции. Таким образом, если модулирующие функции $f_1(t)$ и $f_2(t)$ приводят к появлению боковых спектров φ_1 и φ_2 соответственно, то боковые спектры, получающиеся при модуляции их суммой $f_1(t) + f_2(t)$, будут $\varphi_1 + \varphi_2$. Никаких спектров перекрестной модуляции, которые наблюдались при ЧМ, в данном случае нет. Поэтому амплитудная модуляция, как уже указывалось, является линейной[1].

---

[1]) Общее определение линейной модуляции следующее. Модулированный сигнал есть функция модулирующего сигнала $f(t)$. Пусть $\varphi[f(t)]$ — модулированный сигнал. Модуляция является линейной, если $\{d/d[f(t)]\}\{\varphi[f(t)]\}$ не зависит от $f(t)$. В противном случае она нелинейна. Читатель легко может убедиться в том, что согласно этому определению АМ является линейной, а ЧМ — нет.

Линейная модуляция легко поддается математическим преобразованиям и обобщениям. Спектр модулированного сигнала при модуляции суммой двух функций находится путем вычисления спектров сигнала при модуляции каждой из этих функций и затем суммированием этих спектров. Это заметно облегчает вычисление мощности шумов в системах связи. В случае линейных систем модуляции эффект аддитивного шума в канале можно оценить, предполагая сигнал равным нулю. При нелинейной модуляции так делать нельзя из-за возникновения перекрестной модуляции. По этим причинам представляет интерес аппроксимация нелинейной модуляции некоторой линейной моделью. Это аналогично случаю анализа цепей, когда нелинейная цепь может приближенно рассматриваться как линейная в ограниченном диапазоне значений амплитуд сигнала. Покажем, что при малых значениях индекса модуляции ЧМ хорошо аппроксимируется линейной моделью.

ЛИНЕАРИЗАЦИЯ СИГНАЛА ЧМ ПРИ МАЛЫХ ИНДЕКСАХ МОДУЛЯЦИИ

При малых индексах модуляции ЧМ является почти линейной. Рассмотрим случай, когда модулирующая функция состоит из двух частот ω_1 и ω_2. Сигнал ЧМ при этом записывается в виде ф-лы (4.29). Если m_1 и $m_2 \ll 1$, то

$$\hat{\varphi}_{\text{ЧМ}}(t) \approx A(1 + \mathrm{i}\, m_1 \sin \omega_1 t)(1 + \mathrm{i}\, m_2 \sin \omega_2 t)\, e^{\mathrm{i}\omega_c t} = = A(1 + \mathrm{i}\, m_1 \sin \omega_1 t + \mathrm{i}\, m_2 \sin \omega_2 t)\, e^{\mathrm{i}\omega_c t}. \quad (4.34)$$

Заметим, что если $f_1(t) = a_1 \cos \omega_1 t$, $m_1 \ll 1$, то

$$\hat{\varphi}_{\text{ЧМ}}(t) = A\, e^{\mathrm{i}\, m_1 \sin \omega_1 t}\, e^{\mathrm{i}\omega_c t} \approx A(1 + \mathrm{i}\, m_1 \sin \omega_1 t)\, e^{\mathrm{i}\omega_c t}.$$

Если $f_2(t) = a_2 \cos \omega_2 t$, $m_2 \ll 1$, то

$$\hat{\varphi}_{\text{ЧМ}}(t) \approx A(1 + \mathrm{i}\, m_2 \sin \omega_2 t)\, e^{\mathrm{i}\omega_c t}.$$

Легко видеть, что если m_1, $m_2 \ll 1$, боковые составляющие в случае модуляции $f_1(t) + f_2(t)$ равны сумме боковых составляющих при модуляции функциями $f_1(t)$ и $f_2(t)$ по отдельности. Следовательно, можно считать, что ЧМ является линейной для малых значений индекса модуляции. Слагаемыми перекрестной модуляции при этом допущении можно пренебречь.

4.7. Некоторые замечания о фазовой модуляции

Выше было показано, что при угловой модуляции ширина спектра модулированного сигнала равна приближенно удвоенному наибольшему отклонению частоты. Таким образом, если отклонение частоты поддерживается постоянным, то и спектр модулированного сигнала имеет постоянную ширину.

Для сигналов ЧМ мгновенная частота $\omega_i = \omega_c + ak_f \cos \omega_m t$. Девиация частоты равна ak_f и не зависит от ω_m. Следовательно, для

сигналов ЧМ ширина спектра приближенно равна $2\Delta\omega = 2ak_f$ независимо от частоты модулирующей функции. С другой стороны, для сигналов ФМ:

$$\theta(t) = \omega_c t + ak_p \cos\omega_m t,$$
$$\omega_i(t) = d\theta/dt = \omega_c - ak_p\,\omega_m \sin\omega_m t. \qquad (4.35)$$

Из ф-лы (4.35) вытекает, что девиация частоты при ФМ не является постоянной, а равна $ak_p\omega_m$ и изменяется прямо пропорционально частоте модулирующей функции ω_m. Следовательно, полоса частот, требующаяся для передачи сигналов ФМ, не остается постоянной и сильно зависит от формы модулирующей функции.

На практике, однако, легче получить сигнал ФМ, чем ЧМ. Выше было показано, что если сначала проинтегрировать модулирующую функцию, а затем подать на фазовый модулятор, то получится сигнал ЧМ. Поэтому во многих системах сигналы ЧМ получаются с помощью фазовых модуляторов, на которые в качестве модулирующей функции поступает колебание $\int f(t)dt$.

4.8. Мощности несущей и боковых полос в сигнале угловой модуляции

При рассмотрении сигнала АМ было установлено, что его средняя мощность зависит от коэффициента модуляции m. При угловой модуляции амплитуда сигнала постоянна и не зависит от индекса модуляции m_f. Поэтому разумно ожидать, что мощность сигнала угловой модуляции постоянна и не зависит от интенсивности модулирующей функции. Это действительно так. Обратимся к ф-ле (4.26*а*). Здесь сигнал ЧМ является периодической функцией и выражается как дискретная сумма гармонических слагаемых. Мощность $\varphi_{ЧМ}(t)$ равна сумме мощностей отдельных частотных составляющих[1]. Следовательно,

$$\overline{\varphi^2_{ЧМ}(t)} = \frac{A^2}{2}\sum_{n=-\infty}^{\infty} J_n^2(m_f).$$

Можно показать[2], что сумма в правой части равна единице для всех значений m_f. Поэтому

$$\overline{\varphi^2_{ЧМ}(t)} = A^2/2.$$

Мощность немодулированной несущей $A\cos\omega_c t$ также равна $A^2/2$. Сигнал ЧМ имеет ту же самую мощность, что и немодулиро-

---

[1]) Это следует из теоремы Парсеваля [ф-ла (1.34)].

[2]) Г. Н. Ватсон. Теория бесселевых функций, ИИЛ, 1949.

ванная несущая. Однако модулированный сигнал содержит как несущую, так и боковые полосы, как видно из ф-лы (4.26б). В этой формуле $AJ_0(m_f)$ есть амплитуда частотной составляющей на частоте ω_c, $AJ_n(m_f)$ — амплитуда боковой составляющей n-го порядка. При соответствующем выборе m_f можно обеспечить сколь угодно малую величину $J_0(m_f)$. Действительно, $J_0(m_f)=0$ при $m_f=$ $=2{,}405$; 5,52 и так далее (см. рис. 4.3). Итак, мощность на несущей частоте может быть сделана сколь угодно малой. В этом случае почти вся мощность сигнала переносится боковыми составляющими. Следовательно, при соответствующем выборе m_f эффективность передачи может быть сделана сколь угодно близкой к 100%. Заметим, что при увеличении m_f растет число боковых составляющих и уменьшается $J_0(m_f)$, что и приводит к возрастанию эффективности передачи.

4.9. Подавление шумов при угловой модуляции

Из изложенного выше следует, что при одном и том же модулирующем сообщении требуемая для передачи полоса частот при угловой модуляции значительно шире, чем при АМ. Например, если $f_m=10$ *кгц*, то ширина спектра сигнала АМ равна 20 *кгц*, тогда как ширина спектра сигнала ЧМ составляет около 150 *кгц* при $\Delta f=75$ *кгц*. В случае узкополосной ЧМ ширина спектра сигнала равна ≈ 20 *кгц*. Ни при каких условиях полоса частот, необходимая для передачи сигнала ЧМ, не меньше, чем ширина спектра АМ, что является большим недостатком частотной модуляции. По этой причине частотная модуляция считалась раньше невыгодной и практически мало использовалась. Впоследствии Э. Армстронг показал ошибочность такого взгляда.

Как известно, ЧМ обеспечивает лучшую помехозащищенность при действии шумов и мешающих сигналов. В гл. 7 будет показано, что отношение сигнала к помехе увеличивается на 6 *дб* при увеличении ширины спектра сигнала в два раза. Нетрудно получить снижение шумов до соотношения 30 : 1 *дб* (1000 : 1 по мощности), используя широкополосную ЧМ. Свойство уменьшения шумов при ЧМ следует также непосредственно из теории связи. В гл. 8 показано в самом общем виде, что с увеличением ширины спектра увеличивается и отношение сигнал/шум. С другой стороны, если определено наименьшее допустимое отношение сигнал/шум, то сигнал может передаваться по системе с более узкой полосой пропускания.

Следует напомнить, что свойство подавления шумов при угловой модуляции становится заметным только в том случае, если полоса частот, требуемая для передачи, велика, т. е. если $k_f \gg 1$. Это значит, что индекс модуляции m_f должен быть большим. При узкополосной ЧМ ширина спектра сигнала почти такая же, как при АМ, и улучшение отношения сигнал/шум незначительно.

4.10. Получение сигналов ЧМ

КОСВЕННЫЙ МЕТОД

При косвенном методе получения сигналов ЧМ модулирующее сообщение интегрируется, а затем подается на фазовый модулятор. Этот метод используется довольно часто благодаря простоте получения сигнала ФМ. Из ф-лы (4.8) следует, что если колебание $\int f(t)dt$ подводится ко входу фазового модулятора, то выходной сигнал его будет частотномодулированным. Этот метод уже использовался для получения сигналов узкополосной ЧМ (рис. 4.2). Частотный модулятор, показанный на рис. 4.2б, является, по существу, узкополосным фазовым модулятором, на вход которого поступает сигнал $\int f(t)dt$. Затем можно преобразовать узкополосную ЧМ в широкополосную, используя умножение частоты. Умножитель частоты представляет собой нелинейное устройство, которое умножает частоту входного сигнала. Например, элемент с квадратичной характеристикой может умножить частоту сигнала в два раза. Для такого элемента напряжения на входе $u_i(t)$ и выходе $u_0(t)$ связаны соотношением $u_0(t)=[u_i(t)]^2$. Если $u_i(t)=$ $=\varphi_{\text{ЧМ}}(t)=\cos[\omega_c t+k_f \int f(t)dt]$, то

$$u_o(t)=\cos^2\left[\omega_c t+k_f\int f(t)\,dt\right]=\frac{1}{2}\left\{1+\cos\left[2\omega_c t+2k_f\int f(t)\,dt\right]\right\}.$$

Следовательно, как несущая частота, так и индекс модуляции выходного сигнала увеличились в два раза по сравнению со входными. Вообще, если используется элемент с характеристикой n-й

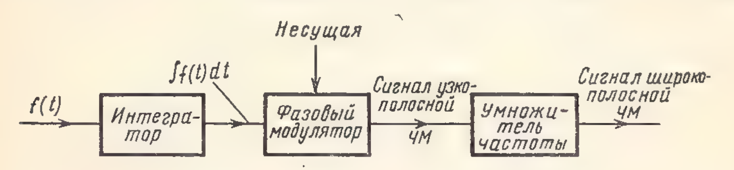


Рис. 4.8. Структурная схема получения сигналов ЧМ косвенным методом

степени, то несущая частота и индекс модуляции умножаются в n раз. Структурная схема получения широкополосной ЧМ с помощью узкополосного фазового модулятора показана на рис. 4.8.

ПРЯМОЙ МЕТОД

При прямом методе получения ЧМ модулирующая функция непосредственно изменяет несущую частоту. Для этой цели используется генератор несущей, у которого параметры одного из реактивных элементов (L или C) изменяются пропорционально $f(t)$.

Если элементы контура генератора обозначаются через L и C, то частота генерации $\omega_i = 1/\sqrt{LC}$.

Когда один из параметров L или C контура изменяется пропорционально $f(t)$, то можно показать, что мгновенная частота генератора также будет изменяться пропорционально $f(t)$ при небольших отклонениях частоты. Допустим, что емкость C контура изменяется по закону:

$$C = C_0 + af(t) = C_0\left[1 + \frac{a}{C_0} f(t)\right]$$

и

$$\omega_i = \frac{1}{\sqrt{LC}\left[1 + \frac{a}{C_0} f(t)\right]^{1/2}}.$$

Если $(a/C_0)f(t) \ll 1$, то

$$\omega_i \approx \frac{1}{\sqrt{LC_0}}\left[1 - \frac{a}{2C_0} f(t)\right] = \omega_c + k_f f(t),$$

где $\omega_c = 1/\sqrt{LC_0}$, $k_f = -a\omega_c/(2C_0)$.

Очевидно, выходное напряжение генератора есть сигнал ЧМ. Подобным же образом можно показать, что и при изменении L получается сигнал ЧМ.

Рассмотрим практические способы получения сигналов ЧМ, реализующие указанный метод.

Модулятор с полупроводниковым диодом. Переменную реактивность можно получить с помощью двухполюсника, сопротивление которого изменяется пропорционально модулирующей функции. Например, емкость запертого полупроводникового диода зависит от подводимого к нему напряжения. Таким образом, подводя к диоду модулирующее напряжение $f(t)$, можно изменять емкость перехода.

Модулятор с насыщенной катушкой индуктивности. В этом модуляторе индуктивность изменяется пропорционально модулирующему напряжению $f(t)$. Проницаемость ферритового сердечника зависит от внешнего магнитного поля, которое можно создать, пропуская ток через дополнительную обмотку, намотанную на сердечник. Через дополнительную обмотку пропускается ток, пропорциональный $f(t)$, что вызывает изменение индуктивности главной катушки пропорционально модулирующему току. Такой способ получения ЧМ весьма распространен благодаря своей простоте, а также потому, что позволяет непосредственно, без умножения частоты, получить девиацию частоты, равную ± 100 *кгц*, требующуюся при ЧМ радиовещании[1]).

---

[1]) Описание других способов получения ЧМ приводится в книге А. Д. Артыма «Теория и методы частотной модуляции». ГЭИ, 1961.—*Прим. ред.*

Модулятор с реактивной лампой. Входное сопротивление реактивной лампы используется как переменное реактивное сопротивление (индуктивное или емкостное), пропорциональное модулирующей функции $f(t)$. Схема реактивной лампы показана на

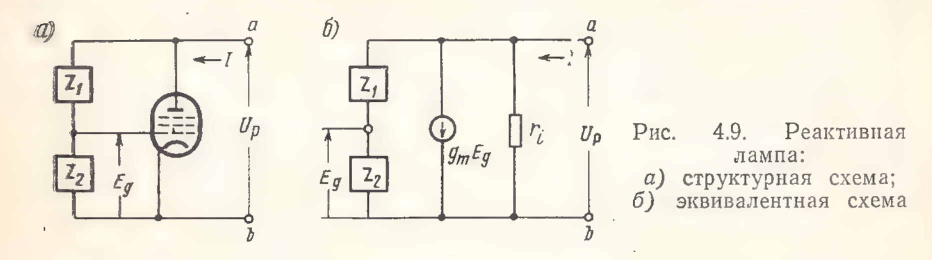


Рис. 4.9. Реактивная лампа:
а) структурная схема; *б)* эквивалентная схема

рис. 4.9*а*, а на рис. 4.9*б* дана ее эквивалентная схема. Поскольку обычно сопротивление r_p велико, его можно не учитывать. Из рис. 4.9*б* получаем

$$I = g_m E_g + \frac{U_p}{Z_1 + Z_2} \tag{4.36}$$

и

$$E_g = \frac{Z_2}{Z_1 + Z_2} U_p. \tag{4.37}$$

Подставим выражение (4.37) в ф-лу (4.36):

$$I = \frac{g_m Z_2 + 1}{Z_1 + Z_2} U_p.$$

Следовательно, проводимость $Y_{ab} = I/U_p = 1 + g_m Z_2/(Z_1 + Z_2)$. Обычно $|g_m Z_2| \gg 1$, и, если $Z_1 \gg Z_2$, то $Y_{ab} \approx g_m Z_2/Z_1$.

Проводимость g_m лампы является функцией напряжения на сетке. Если на сетку подается напряжение, пропорциональное $f(t)$, то

$$g_m = g_{m0} + bf(t), \quad Y_{ab} = [g_{m0} + bf(t)] Z_2/Z_1.$$

Из последнего выражения следует, что если $Z_1 = 1/(i\omega C_1)$ и $Z_2 = R$, то Y_{ab} есть проводимость емкости, изменяющейся во времени:

$$C = C_0 + \beta f(t).$$

В каждом из этих случаев несущая частота и индекс модуляции сохраняются очень малыми. Требуемое значение несущей получается посредством умножения частоты и ее последующего преобразования. При умножении частоты увеличивается также индекс модуляции m_f.

Частотную модуляцию можно также получить с помощью таких приборов, как клистрон и мультивибратор. В этих устройст-

вах частота генерации изменяется при подведении управляющего напряжения в соответствующие точки. В отражательном клистроне частота генерации зависит от напряжения на отражателе. Следовательно, можно модулировать частоту, подводя $f(t)$ к отражателю. В мультивибраторе частота генерации зависит от напряжения на управляющей сетке лампы или базе транзистора.

4.11. Детектирование сигналов ЧМ

Чтобы выделить модулирующее сообщение $f(t)$ из модулированного по частоте сигнала, необходимо иметь цепь, выходное напряжение которой линейно изменяется с изменением частоты входного сигнала. Следовательно, детекторы ЧМ являются частотнозависимыми устройствами и называются также частотными дискриминаторами.

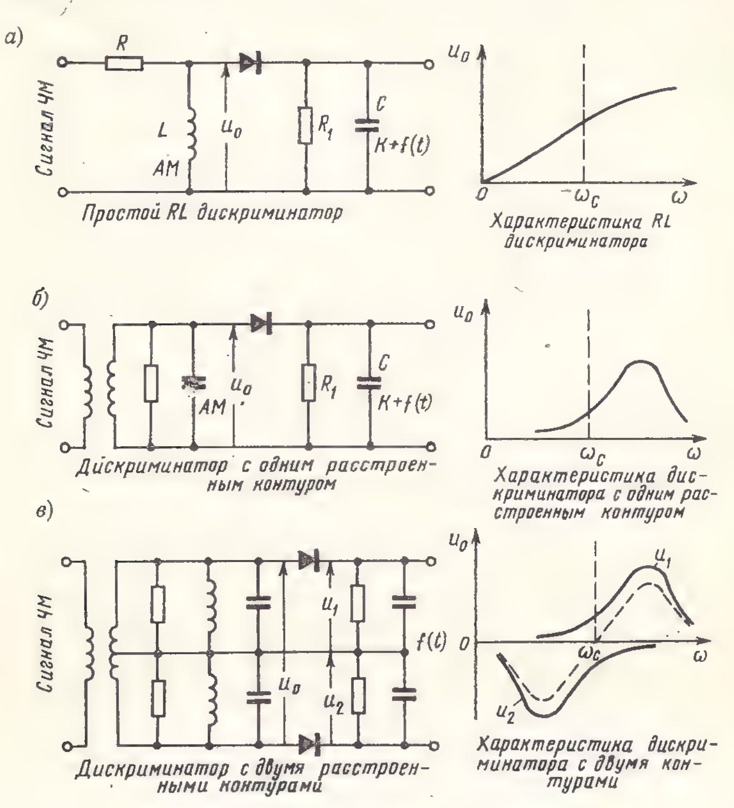


Рис. 4.10. Частотные детекторы:
а) простой *RL*-дискриминатор и его характеристика; *б)* дискриминатор с одним расстроенным контуром и его характеристика; *в)* балансный дискриминатор с двумя расстроенными контурами и его характеристика

Частотный дискриминатор представляет собой цепь, коэффициент передачи которой изменяется линейно с частотой. Таким образом, эта цепь преобразует сигнал ЧМ в сигнал АМ. Далее полученный сигнал АМ детектируется детектором огибающей, состоящим из диода и RC-цепи.

На рис. 4.10 показано несколько простых частотных дискриминаторов. На рис. 4.10*а* цепь RL преобразует сигнал ЧМ в сигнал АМ, который затем детектируется диодом и R_1C-цепью (детектор огибающей). На рис. 4.10*б* сигнал ЧМ преобразуется в сигнал АМ колебательным контуром, который несколько расстроен относительно частоты ω_c. При небольшой девиации частоты напряжение на контуре изменяется почти пропорционально частоте. Получающийся сигнал АМ затем детектируется детектором огибающей. Дискриминатор RL на рис. 4.10*а* имеет низкую чувствительность. Чувствительность дискриминатора с расстроенным колебательным контуром выше, но его характеристика более нелинейна. Балансный дискриминатор, показанный на рис. 4.10*в*, имеет высокую чувствительность и наилучшую линейность. Верхний и нижний контуры его настроены на более высокую и более низкую частоты, чем несущая.

Графики напряжений u_1 и u_2 как функций частоты показаны на рис. 4.10*в*. Выходное напряжение $u_0(\omega)$ показано на этом же рисунке пунктирной линией. Балансный дискриминатор обеспечивает значительно лучшую линейность, чем дискриминатор с одним расстроенным контуром, так как искажения, вызываемые четными гармониками, компенсируются. Кроме того, любые искажения, вызываемые остаточной амплитудной модуляцией, присутствующей в сигнале ЧМ, также уничтожаются.

ЗАДАЧИ

4.1. Покажите, что при угловой модуляции (ЧМ или ФМ) одним гармоническим напряжением $A \cos \omega_m t$ нельзя отличить сигнал ЧМ от сигнала ФМ.

4.2. Несущая 10 *Мгц* модулируется по фазе гармоническим сигналом единичной амплитуды и частоты 10 *кгц*. Наибольшее отклонение фазы равно 10 *рад* на единицу амплитуды модулирующего сигнала. Определите приближенное значение ширины спектра сигнала ФМ. Вычислите ширину спектра сигнала ФМ, если частота модуляции уменьшится до 5 *кгц*. Найдите ширину спектра сигнала ФМ, если частота модулирующего сигнала 10 *кгц*, а амплитуда увеличится в два раза.

4.3. Несущая 100 *Мгц* модулируется по частоте гармоническим сигналом частотой 10 *кгц*, причем девиация частоты равна 1 *Мгц*. Определите ширину спектра сигнала ЧМ. Вычислите ширину спектра сигнала ЧМ, если амплитуда модулирующего сигнала увеличилась в два раза.

Найдите ширину спектра сигнала ЧМ, если также увеличилась в два раза частота модулирующего сигнала.

4.4. Несущая 100 *Мгц* модулируется по фазе гармоническим сигналом частотой 10 *кгц*, причем $k_p = 100$ *рад/в*. Определите ширину спектра сигнала ФМ, если модулирующий сигнал имеет единичную амплитуду. Чему равна ширина спектра сигнала, если амплитуда модулирующей функции увеличится в два раза? Вычислите ширину спектра сигнала, если в два раза возрастет также частота модулирующей функции.

4.5. Несущая модулируется по фазе функцией $a\cos\omega_m t$. Девиация фазы равна k_p [ф-ла (4.8)]. Найдите все возможные соотношения между девиацией частоты $\Delta\omega$, девиацией фазы $\Delta\varphi$, фазовой постоянной k_p, частотой модулирующей функции ω_m и ее амплитудой a.

4.6. Несущая 10 *Мгц* модулируется по фазе гармоническим колебанием частоты 10 *кгц* с единичной амплитудой, причем девиация фазы равна 2 *рад*. Вычислите ширину спектра сигнала ФМ. (Заметим, что девиация частоты в этом случае мала, и для определения ширины спектра нужно использовать график, показанный на рис. 4.5.)

4.7. Сигнал угловой модуляции записывается в виде

$$\varphi(t) = 10\cos(2\cdot 10^6 \pi t + 10\cos 2000\pi t).$$

Найдите: а) мощность модулированного сигнала, б) наибольшее отклонение частоты, в) наибольшее отклонение фазы, г) ширину спектра сигнала.

Можно ли определить, какой сигнал здесь рассматривается: ЧМ или ФМ?

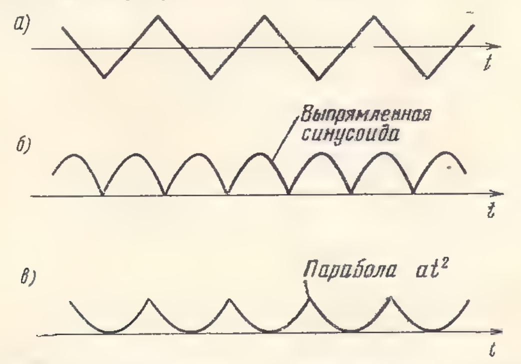
4.8. Несущее колебание модулируется по частоте гармоническим колебанием $f(t)$. Постоянная k_f равна 30 000 *гц/в*. Найдите мощность несущей и боковых составляющих при следующих условиях:

а) $f(t) = \cos 5000t$, г) $f(t) = 2{,}405\cos 30\,000t$,

б) $f(t) = 2\sin 2500t$, д) $f(t) = 10\sin(1000t + \theta)$,

в) $f(t) = \frac{1}{2}\cos 2500t$, е) $f(t) = 5{,}52\sin(30\,000t + \theta)$

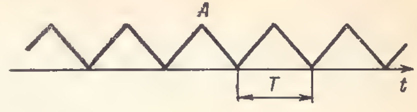
4.9. Нарисуйте осциллограммы сигналов ЧМ и ФМ для модулирующих функций $f(t)$, показанных на рисунке.



4.10 Для каждой из периодических модулирующих функций, рассмотренных в предыдущей задаче, найдите спектр сигнала ЧМ. Определите мощность сигнала на несущей частоте и мощность боковых составляющих.

4.11. Для каждой из периодических модулирующих функций, рассмотренных в задаче 4.9, найдите спектр сигнала ФМ.

4.12. Несущая 100 *Мгц* модулируется по фазе колебанием $f(t)$, показанным на рисунке. Фазовая постоянная модуляции $k_p = 10^6$ *рад/в*. Нарисуйте осциллограмму модулированного сигнала.



Найдите и постройте график спектра сигнала, если:

а) $A = 2\cdot 10^{-6}$, $T = 2\cdot 10^{-6}$ в) $A = 2\cdot 10^{-6}$, $T = 10^{-6}$;

б) $A = 10^{-6}$, $T = 2\cdot 10^{-6}$; г) $A = 10^{-6}$, $T = 10^{-6}$.

ГЛАВА

5

Системы связи с импульсной модуляцией

5.1. Амплитудно-импульсная модуляция

ВВОДНЫЕ ЗАМЕЧАНИЯ

Теорема отсчетов, изложенная в § 1.16, является теоретической основой методов импульсной модуляции. Было показано, что сигнал с ограниченным спектром, который не имеет спектральных составляющих на частотах, выше f_m, *гц*, полностью определяется своими мгновенными значениями, следующими равномерно через интервал $1/(2f_m)$ (или меньше). Поэтому сигнал можно передавать не непрерывно, а лишь в конечное число моментов времени ($2f_m$ раз в секунду). Содержащаяся в каждом отсчете информация может быть передана с помощью импульсной модуляции.

Рассмотрим амплитудно-импульсную модуляцию (АИМ). Здесь значения отсчетов передаются посредством импульсов, амплитуды которых изменяются пропорционально значениям отсчетов. Остановимся сначала на передаче идеальных отсчетов (единичных импульсов), как показано на рис. 5.1. Отсчеты берутся через $1/(2f_m)$, *сек*. Как отмечалось в § 1.15, спектр импульсного сигнала (или выборки) $f_s(t)$ представляет собой периодическое повторение спектра $F(\omega)$ непрерывного сигнала $f(t)$ (рис. 5.1*в*). Из рисунка 5.1*в* следует, что непрерывный сигнал $f(t)$ можно восстановить из импульсного сигнала $f_s(t)$, пропуская $f_s(t)$ через фильтр нижних частот с частотой среза f_m. Спектр импульсного сигнала $f_s(t)$ согласно ф-ле (1.135)

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\,\omega_0) \quad \left(\omega_0 = \frac{2\pi}{T},\ T \leqslant \frac{1}{2f_m}\right).$$

В рассматриваемом случае $T = 1/(2f_m)$. При этом $\omega_0 = 2\,\omega_m$ и

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - 2n\,\omega_m) \left(T = \frac{\pi}{\omega_m}\right). \tag{5.1}$$

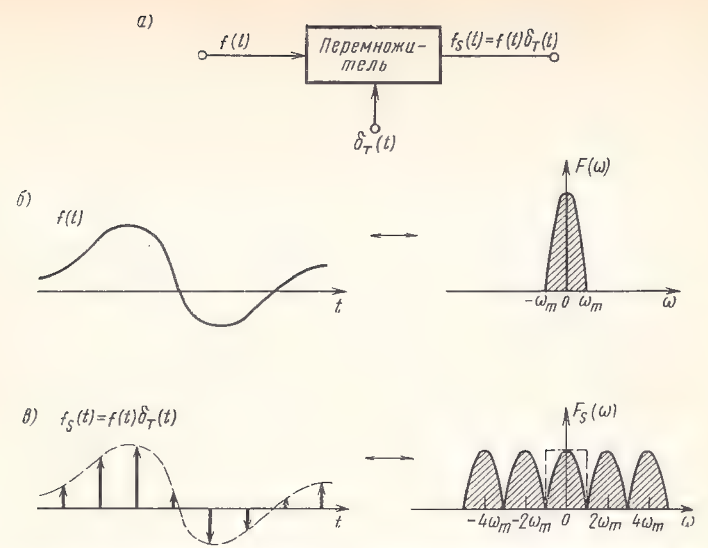


Рис. 5.1. К пояснению АИМ с использованием единичных импульсов:
а) модулятор-перемножитель; *б)* модулирующий сигнал и его спектр; *в)* сигнал АИМ (выборка) и его спектр

ЕСТЕСТВЕННАЯ ВЫБОРКА

В рассмотренном процессе отсчеты функции $f(t)$ образовывались с помощью единичных импульсов (мгновенная выборка). Из рис. 5.1*в* видно, что спектр такого идеального импульсного сигнала занимает всю бесконечную полосу частот от $-\infty$ до $+\infty$, т. е. содержит составляющие всех частот. Однако на практике такую идеальную выборку осуществить нельзя, так как невозможно генерировать единичные импульсы. Обычно выборка осуществляется с помощью очень узких импульсов конечной длительности. Следовательно, в этом случае отсчеты не являются мгновенными, а занимают некоторый конечный временной интервал. Рассмотрим свойства такого импульсного сигнала.

Допустим, что выборка выполняется с помощью периодической последовательности импульсов прямоугольной формы, имеющих длительность τ *сек* и следующих через каждые $T=1/(2f_m)$ *сек*. Обозначим эту последовательность импульсов $p_\tau(t)$. Спектр $P_\tau(\omega)$ сигнала $p_\tau(t)$ показан на рис. 5.2*б* [см. также ф-лу (1.103) и рис. 1.23]. Импульсный сигнал $f_s(t)$ есть произведение $f(t)$ на $p_\tau(t)$. Следовательно, его спектр $F_s(\omega)$ получается сверткой спектра $F(\omega)$ со спектром $P_\tau(\omega)$. Свертка легко выполняется графически и ее результат показан на рис. 5.2*в*. При неидеальной выборке функции $f(t)$ образуется спектр, подобный спектру при идеальной выборке, но с убывающими амплитудами. Этот же результат

можно получить аналитически. В рассматриваемом случае $f_s(t) = f(t)p_\tau(t)$. Следовательно,

$$F_s(\omega) = \frac{1}{2\pi} F(\omega) * P_\tau(\omega). \quad (5.2)$$

Спектр $P_\tau(\omega)$ можно найти из ф-лы (1.103), подставив в нее $T = 1/(2f_m) = \pi/\omega_m$ и $\omega_0 = 2\pi/T = 2\omega_m$:

$$P_\tau(\omega) = 2A\tau\omega_m \sum_{n=-\infty}^{\infty} Sa(\pi\tau\omega_m)\,\delta(\omega - 2n\omega_m). \quad (5.3)$$

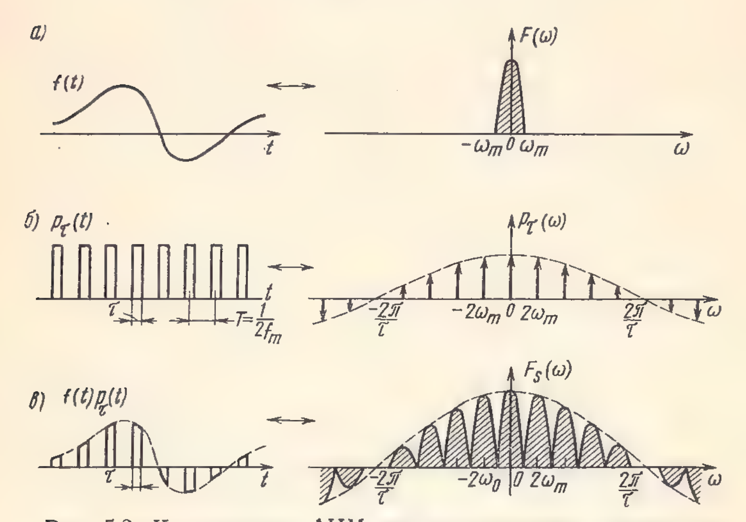


Рис. 5.2. К пояснению АИМ при использовании прямоугольных импульсов конечной длительности:
а) модулирующий сигнал и его спектр; *б)* периодическая последовательность прямоугольных импульсов и ее спектр; *в)* сигнал АИМ и его спектр

Подставив ф-лу (5.3) в (5.2), получим

$$F_s(\omega) = \frac{A\tau\omega_m}{\pi} F(\omega) * \sum_{n=-\infty}^{\infty} Sa(n\tau\omega_m)\,\delta(\omega - 2n\omega_m) =$$

$$= \frac{A\tau}{T} \sum_{n=-\infty}^{\infty} Sa(n\tau\omega_m)\,F(\omega) * \delta(\omega - 2n\omega_m) =$$

$$= \frac{A\tau}{T} \sum_{n=-\infty}^{\infty} Sa(n\tau\omega_m)\,F(\omega - 2n\omega_m). \quad (5.4)$$

Правая часть ф-лы (5.4) представляет спектр $F(\omega)$, повторяющийся через каждые $2\,\omega_m$, *рад/сек*, причем амплитуды спектра изменяются как функции $Sa(n\tau\omega_m)$. Этому выражению соответствует спектр, показанный на рис. 5.2*в*. Заметим, что импульсы не обязательно должны быть прямоугольными, как на рис. 5.2*в*. Отсчеты могут представляться с помощью импульсов любой другой формы $q(t)$. При этом изменится лишь огибающая спектра $F_s(\omega)$ на рис. 5.2*в*. [Эта огибающая $Q(\omega)$ находится как преобразование Фурье функции $q(t)$].

Итак, при неидеальной выборке сигнала $f(t)$ получается периодически повторяющийся спектр $F(\omega)$, но с убывающими амплитудами. Исходный сигнал $f(t)$ можно выделить из импульсного сигнала $f_s(t)$, используя фильтр нижних частот с частотой среза ω_m. Процесс демодуляции при импульсной модуляции (неидеальной выборке) является тем же самым, что и при идеальной выборке с помощью единичных импульсов. Заметим, что сигнал восстанавливается без искажений, несмотря на то, что выборка неидеальная. Для передачи сигнала при идеальной выборке требуется бесконечно широкая полоса частот, тогда как при неидеальной выборке (импульсной модуляции) требуется конечная полоса частот, поскольку спектр $F_s(\omega)$ (рис. 5.2*в*) убывает с ростом частоты и на очень высоких частотах содержится пренебрежимо малая энергия.

По мере увеличения ширины импульсов спектр убывает быстрее и, следовательно, для передачи требуется меньшая полоса частот. В этом отношении неидеальная выборка превосходит идеальную, поскольку требует меньшую полосу частот для передачи. Однако то, что выигрывается в частотной области, проигрывается во временно́й. Импульсная модуляция при конечной длительности импульсов занимает большее время, чем мгновенная передача в случае единичных импульсов. Поскольку импульсы имеют конечную длительность, можно передавать лишь конечное число сигналов одновременно при временно́м разделении (это называется временны́м уплотнением), тогда как в случае единичных импульсов можно передавать любое число сигналов.

Рассмотренный импульсный сигнал можно выразить как произведение $f(t)$ на периодическую последовательность импульсов $q\ (t)$:

$$f_s(t) = f(t)\, q_\tau(t) = f(t) \sum_{n=-\infty}^{\infty} q(t - nT),$$

где $q(t)$ — импульс последовательности.

Выборка такого типа называется естественной [1]).

---

[1]) В отечественной литературе этот вид модуляции называется амплитудно-импульсной модуляцией первого рода. См. Фельдбаум и др. Теоретические основы связи и управления. Физматгиз, 1963. — *Прим. ред.*

При естественной выборке каждый импульс умножается на $f(t)$ в пределах его длительности. В результате каждый импульс в сигнале $f_s(t)$ имеет разную форму. Это хорошо видно из рис. 5.2*в*, где вершина каждого импульса принимает форму модулирующей функции $f(t)$ в пределах длительности импульса. Теперь рассмотрим мгновенную выборку[1]), при которой все импульсы в сигнале $f_s(t)$ имеют одну и ту же форму, а их амплитуды пропорциональны значениям соответствующих отсчетов (рис. 5.3). Очевидно, такой им-

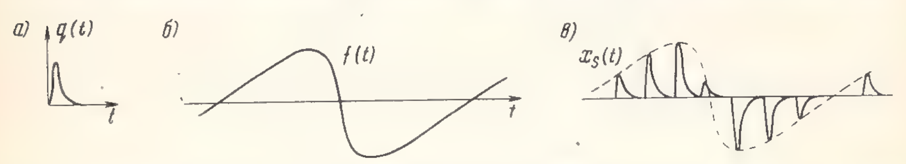


Рис. 5.3. К пояснению мгновенной выборки:
а) одиночный импульс последовательности; *б)* модулирующий сигнал; *в)* сигнал АИМ

пульсный сигнал несет полную информацию, содержащуюся во всех отсчетах, и, следовательно, содержит всю информацию, относящуюся к непрерывной функции $f(t)$ [при условии, что интервал выборки равен или меньше $1/(2f_m)$, *сек*]. Заметим, что при естественной выборке (рис. 5.2) информация о $f(t)$ переносится в течение всей длительности каждого импульса. Напротив, сигнал при мгновенной выборке содержит информацию о $f(t)$ только в моменты отсчетов. Именно поэтому такая выборка называется мгновенной. Следует заметить, что выборка с помощью единичных импульсов (рис. 5.1) может рассматриваться и как естественная, и как мгновенная.

Пусть одиночный импульс из последовательности, используемой при мгновенной выборке, имеет форму $q(t)$, как показано на рис. 5.3*а*. Импульсный сигнал обозначим через $x_s(t)$. В случае мгновенной выборки

$$x_s(t) = \sum_{n=-\infty}^{\infty} f(nT)\, q(t-nT). \tag{5.5}$$

Заметим, что обозначение $f_s(t)$ резервируется для случая, когда выборка $f(t)$ осуществляется единичными импульсами:

$$f_s(t) = \sum_{n=-\infty}^{\infty} f(nT)\, \delta(t-nT). \tag{5.6}$$

[1]) Мгновенная выборка называется также амплитудно-импульсной модуляцией второго рода. — *Прим. ред.*

Определим спектр сигнала $x_s(t)$ при произвольной форме импульса $q(t)$. Он легко находится, если заметить, что $x_s(t)$ есть отклик цепи, имеющей импульсную реакцию $q(t)$, на входной сигнал $f_s(t)$ (рис. 5.4). Таким образом, если $q(t) \longleftrightarrow Q(\omega)$, то $x_s(t) \longleftrightarrow F_s(\omega)Q(\omega)$. Воспользовавшись ф-лой (5.1), получим

$$x_s(t) \longleftrightarrow \frac{1}{T} \sum_{n=-\infty}^{\infty} Q(\omega) F(\omega - 2n\,\omega_m) \left(T = \frac{\pi}{\omega_m}\right). \qquad (5.7)$$

Из ф-лы (5.7) следует, что спектр импульсного сигнала $x_s(t)$ состоит из периодически повторяющихся спектров $F(\omega)$, каждый из которых умножается на спектр единичного импульса $Q(\omega)$. Этот спектр для случая прямоугольных импульсов показан на рис. 5.5г.

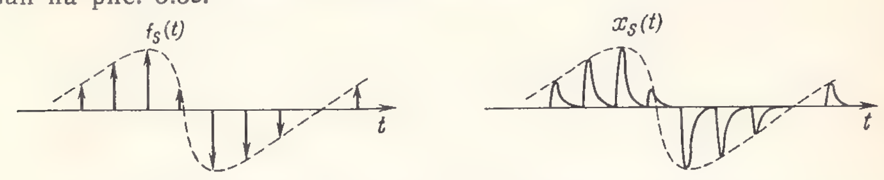


Рис. 5.4. Представление сигнала при мгновенной выборке как отклика цепи с импульсной реакцией $h(t)=g(t)$ на последовательность модулированных единичных импульсов

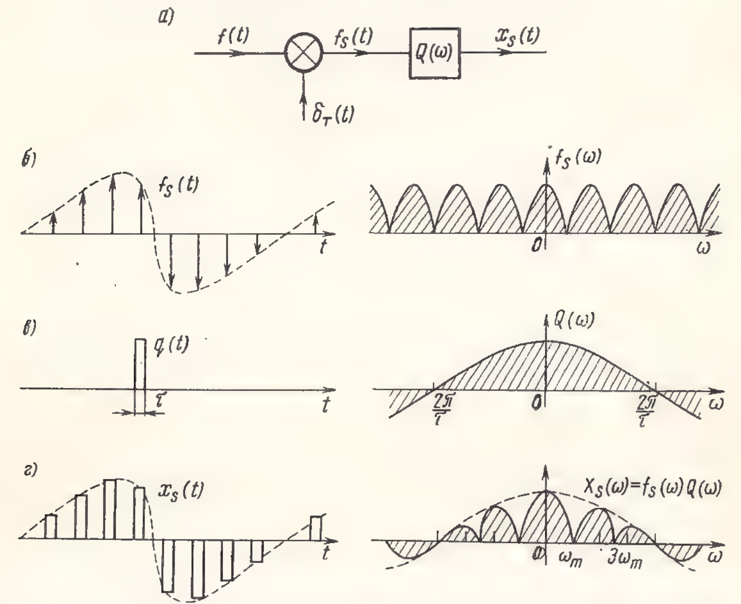


Рис. 5.5. К определению спектра сигнала при мгновенной выборке: *а)* модулятор; *б)* модулированные единичные импульсы и их спектр; *в)* одиночный прямоугольный импульс и его спектр; *г)* сигнал АИМ и его спектр

Заметим, что $X_s(\omega)$ [спектр сигнала $x_s(t)$] на рис. 5.5г не такой, как на рис. 5.2в, хотя на первый взгляд они кажутся одинаковыми. На рис. 5.2в спектр состоит из периодически повторяющихся спектров $F(\omega)$ с убывающими амплитудами. Однако в каждом периоде форма $F(\omega)$ сохраняется неизменной. Изменяются только амплитуды. С другой стороны, на рис. 5.5г форма спектра в каждом интервале $2\omega_m$ отличается от $F(\omega)$ вследствие умножения на $Q(\omega)$. Этот множитель имеет разные значения на всех частотах. Таким образом, ни один из спектров на рис. 5.5г не имеет формы спектра $F(\omega)$. Это следует также из ф-л (5.4) и (5.7). В ф-ле (5.4) множитель $Sa(n\tau\omega_m)$ постоянен для данного n, тогда как в ф-ле (5.7) множитель $Q(\omega)$ является функцией частоты.

ВОССТАНОВЛЕНИЕ $f(t)$ ПРИ МГНОВЕННОЙ ВЫБОРКЕ

Спектр импульсного сигнала $x_s(t)$ (рис. 5.5г) не содержит неискаженный спектр $F(\omega)$, как это было в случае естественной выборки (рис. 5.2в). Таким образом, используя только идеальный фильтр нижних частот, невозможно восстановить сигнал $f(t)$ без искажений.

Если используется фильтр нижних частот с частотой среза ω_m, то спектр на выходе будет $F(\omega)Q(\omega)$. Неискаженный сигнал $f(t)$ можно восстановить из этого выходного напряжения, пропустив его через другой фильтр, который имеет передаточную функцию $1/Q(\omega)$ (рис. 5.6а). Заметим, что поскольку спектр $F(\omega)$ ограничен частотой ω_m, то достаточно, чтобы второй фильтр на рис. 5.6а имел передаточную функцию $1/Q(\omega)$ только в интервале частот $(-\omega_m, \omega_m)$. За пределами этой полосы частот передаточная функция может выбираться произвольно в зависимости от удобства реализации. Оба фильтра можно объединить в один составной фильтр, частотная характеристика которого показана на рис. 5.6б. Очевидно, передаточная функция $H(\omega)$ этого фильтра должна быть равна $1/Q(\omega)$ в интервале частот $(-\omega_m, \omega_m)$ и равна нулю

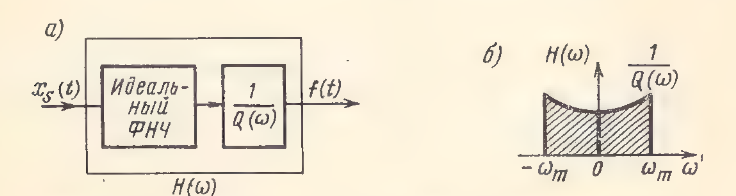


Рис. 5.6. К пояснению восстановления непрерывного сигнала при мгновенной выборке:
а) фильтр нижних частот, составленный из идеального *ФНЧ* и фильтра с передаточной функцией $1/Q(\omega)$;
б) передаточная функция *ФНЧ*

на всех других частотах. Таким образом, передаточную функцию фильтра, восстанавливающего непрерывный сигнал, можно выразить в виде

$$H(\omega) = \begin{cases} \dfrac{1}{Q(\omega)} & \text{при } |\omega| < \omega_m, \\ 0 & \text{при } |\omega| > \omega_m. \end{cases} \tag{5.8}$$

Если импульс $q(t)$ является очень узким, то он приближается к единичному, а спектр $Q(\omega)$ стремится к равномерному. Иными словами, спектр $Q(\omega)$ становится почти постоянным на интервале $(0, \omega_m)$. В таком случае составной фильтр, показанный на рис. 5.6, превращается в идеальный фильтр нижних частот с частотой среза ω_m, и сигнал $f(t)$ восстанавливается так же, как при естественной выборке. Этого следовало ожидать, так как, если импульсы становятся очень узкими, они приближаются к единичным и выборка приближается к выборке с помощью единичных импульсов.

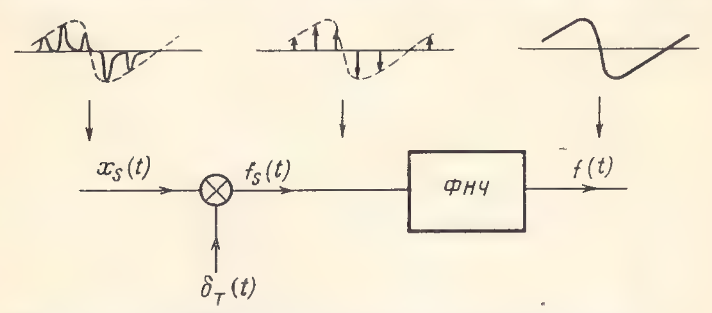


Рис. 5.7. К пояснению восстановления непрерывного сигнала умножением на последовательность единичных импульсов с последующей фильтрацией идеальным *ФНЧ*

С другой стороны, можно восстановить непрерывный сигнал $f(t)$ из импульсного $x_s(t)$, если умножить $x_s(t)$ на последовательность единичных импульсов $\delta_T(t)$. Это, очевидно, дает сигнал $f_s(t)$, как показано на рис. 5.7. Теперь сигнал $f(t)$ можно выделить из $f_s(t)$, используя идеальный фильтр нижних частот.

ПЕРЕДАЧА СИГНАЛОВ АИМ

Сигналы АИМ могут непосредственно передаваться по линии, состоящей из пары проводов. Но передавать их непосредственно в свободном пространстве с помощью электромагнитных волн не так легко, поскольку спектр сигнала АИМ сконцентрирован на низких частотах, для излучения которых требуются антенны практически невыполнимых размеров. Поэтому спектр перемещается на высокие частоты посредством методов амплитудной модуляции, кото-

рые рассматривались в гл. 3. В приемнике сигнал подвергается демодуляции, чтобы осуществить обратный перенос спектра в его первоначальное положение. На выходе этого демодулятора образуется последовательность импульсов $f_s(t)$, из которой с помощью фильтра нижних частот выделяется непрерывный сигнал $f(t)$. Такие системы обозначаются как АИМ/АМ.

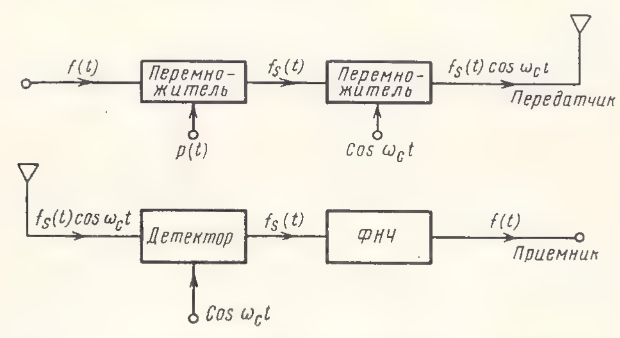


Рис. 5.8. Структурные схемы передатчика и приемника системы связи с АИМ/АМ

Перенос спектра сигнала АИМ на высокие частоты можно также осуществить посредством угловой модуляции. Такие системы обозначаются как АИМ/ЧМ или АИМ/ФМ. Структурная схема системы АИМ/АМ показана на рис. 5.8.

5.2. Дяугие виды импульсной модуляции

Чтобы передать сигнал $f(t)$, необходимо передать информацию о его мгновенных значениях, следующих друг за другом через интервалы времени $1/(2f_m)$, *сек.* Существенным в этом утверждении является то, что нет необходимости в непрерывной передаче сигнала с ограниченным спектром, а полная информация о таких сигналах может быть передана их значениями в дискретные моменты времени. Такая дискретная форма информации может передаваться многими способами. Рассмотренная в § 5.1 амплитудно-импульсная модуляция не является единственным примером.

В системах АИМ эта информация переносится значениями амплитуд импульсов. Можно было бы поддерживать амплитуды всех импульсов постоянными, а изменять их ширину пропорционально значениям $f(t)$ в соответствующие моменты времени. Такой вид модуляции называют широтно-импульсной модуляцией (ШИМ), или модуляцией по длительности импульсов (ДИМ). Еще один способ передачи заключается в том, что как амплитуда, так и ширина импульсов сохраняются неизменными, а изменяется временное положение импульсов пропорционально значениям отсчетов $f(t)$ в соответствующие моменты времени. Такие системы изве-

стны как системы фазово-импульсной модуляции (ФИМ). Сигналы АИМ, ШИМ и ФИМ показаны на рис. 5.9.

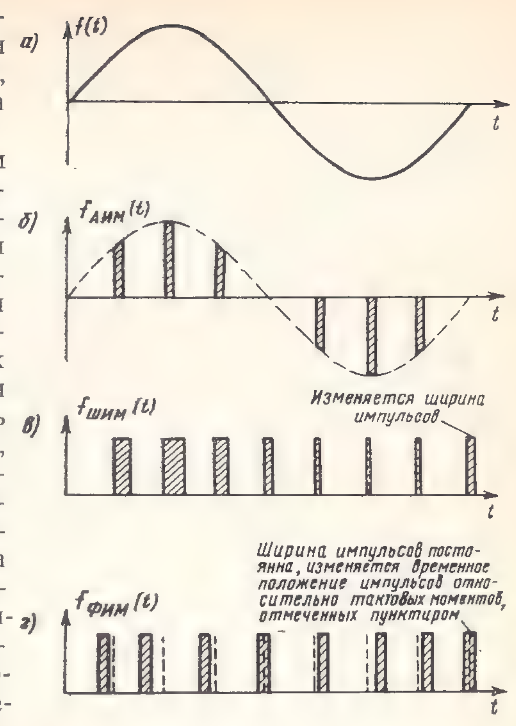


Рис. 5.9. Различные виды импульсной модуляции:
а) модулирующий сигнал; *б)* сигнал АИМ; *в)* сигнал ШИМ; *г)* сигнал ФИМ

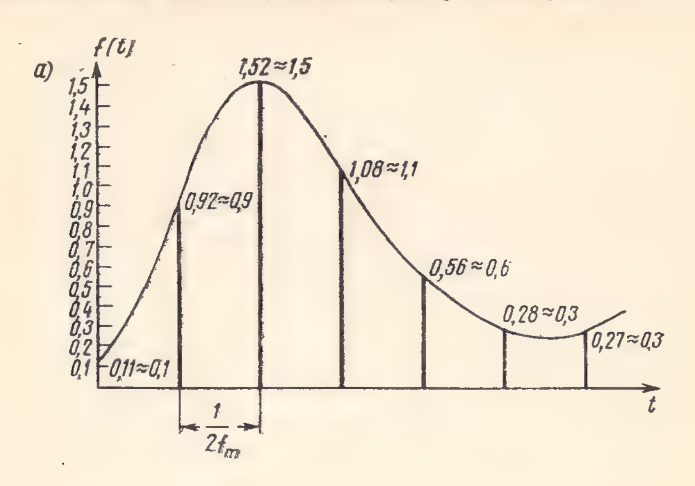
Другим очень важным способом импульсной модуляции является кодово-импульсная модуляция (КИМ). В этой системе каждый отсчет $f(t)$ передается кодовой комбинацией, составленной из нескольких импульсов. Из сообщения $f(t)$, которое должно быть передано, берется выборка, затем каждый отсчет округляется до значения ближайшего разрешенного уровня[1]), как показано на рис. 5.10*а*. На этом рисунке имеется 16 разрешенных уровней, расположенных через 0,1 *в*. Каждый уровень представляется определенной комбинацией импульсов. Один из таких кодов показан на рис. 5.10*б*. Таким образом, вместо отдельных отсчетов передается соответствующая кодовая комбинация.

В примере на рис. 5.10*б* для образования кодовых комбинаций используются двоичные импульсы (импульсы, которые могут принимать два значения). В общем случае могут использоваться s-ичные импульсы (импульсы, которые могут принимать s значений). Чем больше значение s, тем меньше требуется разрядов в кодовой комбинации. Таким образом, при КИМ вместо одного импульса на отсчет (как в случае АИМ) необходимо передавать группу импульсов. Поэтому КИМ требует большего времени[2]) для передачи одинаковой информации, чем другие способы модуляции. Однако этот недостаток КИМ перекрывается ее большей помехоустойчивостью. Для приема сигнала КИМ требуется толь-

[1]) Процесс замены мгновенных значений отсчетов ближайшими разрешенными называется квантованием сигнала. — *Прим. ред.*

[2]) Или более широкой полосы частот, если информация передается в реальном масштабе времени. — *Прим. ред.*

ко установить, присутствует импульс или отсутствует независимо от его амплитуды и формы. Таким образом, внешняя помеха, которая вносит искажение высоты или ширины импульсов, не влияет на прием сигналов КИМ. (Этот вопрос рассматривается в гл. 7.)



б)

Дискретное сообщение	Двоичное числа	Сигнал двоичной КИМ
0	0000	
1	0001	
2	0010	
3	0011	
4	0100	
5	0101	
6	0110	
7	0111	
8	1000	
9	1001	
10	1010	
11	1011	
12	1100	
13	1101	
14	1110	
15	1111	

Рис. 5.10. К пояснению преобразования непрерывного сигнала при кодово-импульсной модуляции:
а) округление отсчетов до ближайших разрешенных значений (квантование); *б)* комбинации четырехразрядного двоичного кода, соответствующие квантованным значениям сигнала

ПЕРЕДАТЧИК И ПРИЕМНИК СИСТЕМЫ С ВРЕМЕННЫМ УПЛОТНЕНИЕМ

На основании теоремы отсчетов можно передавать всю информацию, содержащуюся в непрерывном сигнале с ограниченным спектром, в виде отсчетов этого сигнала через равные интервалы времени. Для передачи отсчетов канал используется не все время, и поэтому можно, используя временно́е разделение, одновременно передавать несколько сигналов. Для этого образуются выборки всех подлежащих передаче сигналов, которые затем передаются в интервалах между импульсами других сигналов, как показано на рис. 5.11. В приемнике отсчеты, принадлежащие каждому сигналу, выделяются с помощью соответствующих устройств.

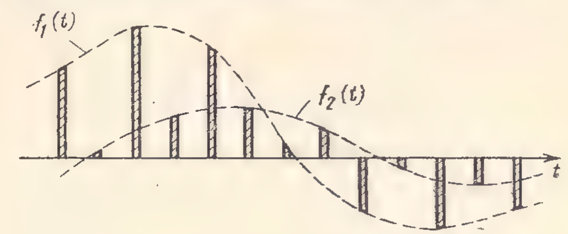


Рис. 5.11. Временно́е уплотнение двух сигналов

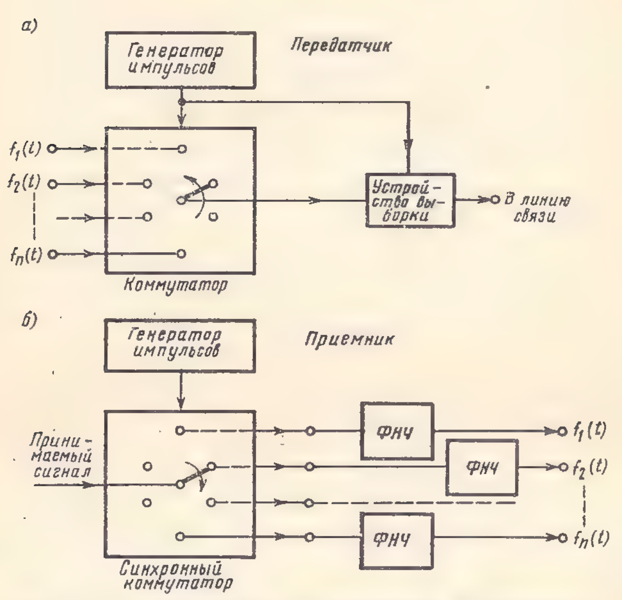


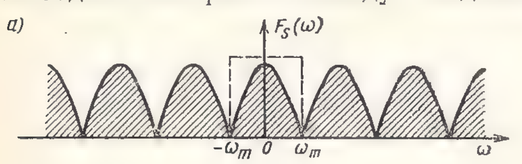
Рис. 5.12. Система связи с временным уплотнением: *а)* передатчик; *б)* приемник

Кратко опишем передатчик и приемник системы с временным уплотнением (рис. 5.12). В передатчике коммутатор, управляемый сигналом от генератора импульсов, последовательно подключается ко всем источникам сообщений. Этот же задающий генератор импульсов вырабатывает отсчетные импульсы. Таким образом, коммутатор последовательно подключает источник сообщений к устройству выборки, которое последовательно берет отсчеты с помощью импульсов, вырабатываемых задающим генератором. Коммутатор и устройство выборки действуют синхронно. Таким образом, выходом устройства выборки является сигнал, который состоит из последовательных отсчетов всех сообщений. В приемнике для распределения импульсов по отдельным каналам используется свой задающий импульсный генератор и коммутатор, который действует синхронно с коммутатором передатчика.

Отсчеты различных сообщений оказываются разделенными. Непрерывное сообщение в каждом канале восстанавливается с помощью фильтров нижних частот.

ЧАСТОТА ВЫБОРКИ

Для того чтобы отсчеты сообщения с ограниченным спектром несли всю информацию об этом сообщении, частота выборки должна быть не менее $2f_m$ отсчетов в секунду. Это минимальная частота отсчетов. Очевидно, при такой частоте отсчетов спектр сообщения повторяется по частоте без перекрытия и без каких-либо свободных интервалов между соседними периодами, как показано на рис. 5.13*а*. Для восстановления непрерывного сообщения $f(t)$ из импульсного сигнала необходимо использовать идеальный фильтр нижних частот, который пропускает все частоты $\omega<\omega_m$ и задерживает все частоты выше ω_m. Практический фильтр нижних частот с резким спадом характеристики можно построить лишь при очень большом числе элементов. Можно ослабить требования к фильтру, если осуществлять выборку сообщения с более высокой, чем минимальная, частотой. Легко видеть, что при этом спектр сообщения повторяется со свободными частотными промежутками (или *защитными полосами*), как показано на рис. 5.13*б*. Из рисунка



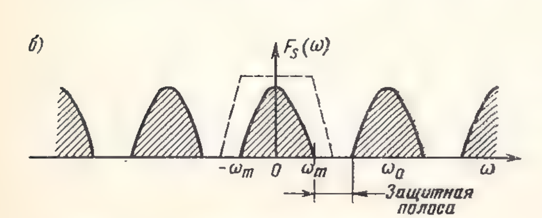


Рис. 5.13. Спектры сигналов АИМ: *а)* при минимальной частоте отсчетов $1/(2f_m)$; *б)* при частоте отсчетов, превышающей минимальную

следует, что непрерывное сообщение можно восстановить фильтром нижних частот, характеристика которого не должна иметь резкого спада.

Подобное же положение возникает при частотном уплотнении каналов. Спектры различных сигналов разделяются защитными полосами по вышеупомянутым причинам. При временно́м уплотнении отсчеты различных сообщений передаются в свободные интервалы времени между импульсами других каналов. Если импульсы расположены очень близко друг от друга, то это вызывает повышенные требования к системе разделения отсчетов разных каналов в приемнике. Поэтому отсчеты последовательных сообщений также разделяются небольшими временны́ми интервалами, которые называются *защитными временны́ми интервалами.*

5.4. Полоса частот, необходимая для передачи сигналов АИМ

Для сравнения АИМ с другими видами модуляции важно знать полосу частот, необходимую для передачи сигналов АИМ. Спектр сигнала АИМ, как видно из рис. 5.1*в*, 5.2*в* или 5.5*г*, содержит составляющие во всем бесконечном интервале частот $(-\infty, \infty)$. Поэтому для совершенно точной передачи сигнала АИМ требуется бесконечная полоса частот. Покажем, однако, что вся содержащаяся в отсчетах информация может быть полностью передана в ограниченной полосе частот.

Если спектр каждого сообщения ограничен частотой f_m, *гц*, то при временно́м уплотнении n непрерывных сообщений необходимо передавать $2f_m$ отсчетов в секунду на одно сообщение. Чтобы передать n сообщений, частота импульсов должна быть $2nf_m$, *имп/сек*. Покажем, что информация, содержащаяся в этих $2nf_m$ импульсах, может быть передана в полосе частот nf_m, *гц*.

На первый взгляд может показаться, что для передачи сигналов АИМ необходима бесконечная полоса частот, потому что передаются импульсы прямоугольной формы, спектры которых занимают всю бесконечную область частот. Однако следует помнить, что в действительности форма импульсов интереса не представляет. Необходимо знать лишь их амплитуды. Поэтому допустимы любые искажения формы импульсов, при которых их амплитуды сохраняются неизменными.

Выходной сигнал устройства временно́го уплотнения состоит из $2nf_m$ импульсов в секунду, причем импульсы каждого из n сообщений повторяются с частотой $2f_m$ отсчетов в секунду. Можно определить полосу частот, которая требуется для передачи $2nf_m$ независимых отсчетов в секунду. Из теоремы отсчетов следует, что непрерывное сообщение со спектром, ограниченным частотой B, *гц*, может быть передано с помощью $2B$ независимых отсчетов в се-

кунду[1]). Наоборот, можно утверждать, что $2nf_m$ независимых отсчетов в секунду определяют непрерывный сигнал, ограниченный по спектру частотой nf_m, *гц*. Можно рассматривать эти $2nf_m$ отсчетов как отсчеты некоторого непрерывного сигнала $\varphi(t)$, спектр которого ограничен частотой nf_m, *гц*. Действительно по ф-ле (1.140) можно получить этот сигнал $\varphi(t)$ из отсчетов. Поэтому вместо передачи $2nf_m$ отсчетов в секунду можно передавать соответствующий непрерывный сигнал $\varphi(t)$, определяемый этими отсчетами. Поскольку спектр указанного сигнала ограничен nf_m, *гц*, то для его передачи требуется полоса частот nf_m, *гц*.

Как же можно получить непрерывный сигнал $\varphi(t)$ из $2nf_m$ отсчетов в секунду? Дискретные импульсы являются просто отсчетами $\varphi(t)$, и поэтому из предыдущего рассмотрения следует, что $\varphi(t)$ можно получить, пропустив отсчеты через фильтр нижних частот с частотой среза nf_m, *гц*.

Обозначим сигнал, содержащий $2nf_m$, *имп/сек*, через $\varphi_s(t)$, а его спектр — через $\Phi_s(\omega)$. Заметим, что $\Phi_s(\omega)$ образуется периодическим повторением спектра $\Phi(\omega)$. Сигналы $\varphi(t)$ и $\varphi_s(t)$ и их спектры показаны на рис. 5.14. Этот результат очень важен, и он

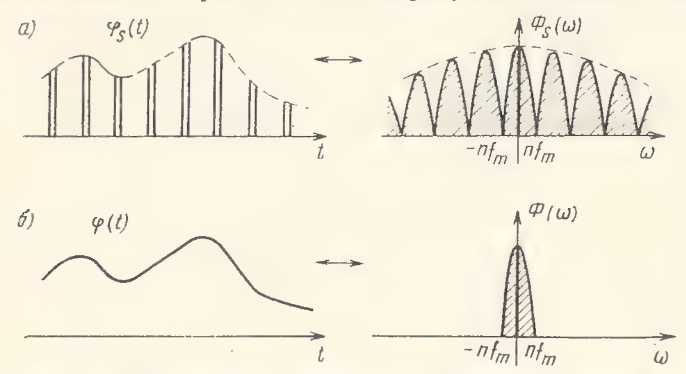


Рис. 5.14. К пояснению фильтрации сигнала $\varphi(t)$ из многоканального импульсного сигнала $\varphi_s(t)$:
а) сигнал $\varphi_s(t)$ и его спектр; б) сигнал на выходе идеального *ФНЧ* с частотой среза nf_m и его спектр

еще будет использоваться в последующих главах. Однако следует напомнить, что если требуется восстановить не только амплитуду, но и форму импульса, то необходима значительно более широкая полоса частот, чем nf_m, *гц*.

Интересно отметить, что полоса частот, необходимая для непосредственной передачи n сигналов АИМ с временны́м разделением,

[1]) Заметим, что если отсчетов в секунду больше, чем $2B$, то они не являются независимыми. Максимум $2B$ отсчетов являются независимыми, а остальные отсчеты выражаются через $2B$ независимых отсчетов. Это легко увидеть из ф-лы (1.140). Вся информация о сигнале может быть выражена минимум $2B$ отсчетами в секунду. Поэтому остальные отсчеты всегда можно выразить через эти $2B$ отсчетов.

в точности равна полосе частот, которая требуется для передачи этих n сигналов с использованием АМ-ОБП и частотного уплотнения. Однако, как уже отмечалось, непосредственная передача сигналов АИМ посредством излучения практически неосуществима, поскольку энергия таких сигналов сконцентрирована на низких частотах, а их излучение требует антенн очень большого размера. В этих случаях весь спектр сигналов АИМ сдвигается на высокие частоты посредством амплитудной модуляции.

При амплитудной модуляции появляются верхняя и нижняя боковые полосы, а значит, необходимая для передачи полоса частот увеличивается в два раза. Таким образом, полоса частот, необходимая для передачи сигнала АИМ/АМ при временно́м уплотнении n непрерывных сообщений, спектры которых ограничены частотой f_m, равна $2nf_m$. Следует заметить, что необходимая для передачи этих n сообщений полоса частот при использовании АМ-ДБП и частотном уплотнении также равна $2nf_m$, *гц*. Поэтому можно сделать вывод, что для передачи сигнала АИМ требуется та же полоса частот, что и для передачи АМ-ОБП. Подобным образом, для передачи сигнала АИМ/АМ необходима такая же полоса частот, как для передачи АМ-ДБП.

5. 5. Сравнение систем частотного и временно́го уплотнения

Выше были рассмотрены два способа одновременной передачи нескольких сообщений с ограниченными спектрами. В системах с частотным уплотнением все сигналы, которые должны быть переданы, являются непрерывными и смешиваются во временно́й области. Однако спектры различных модулированных сигналов занимают разные полосы в частотной области и могут быть разделены с помощью соответствующих фильтров.

С другой стороны, при временно́м уплотнении отсчеты каждого сигнала остаются различными и могут быть разделены во временно́й области. Однако частотные спектры этих различных выборочных сигналов занимают одну и ту же полосу частот. Итак, при частотном уплотнении сохраняется частотный признак сигнала, тогда как признак формы колебания сохраняется при временно́м уплотнении. Поскольку сигнал полностью определяется либо своим частотным представлением, либо временны́м, уплотняемые сигналы можно разделить в приемнике соответствующими способами.

Различие между двумя способами уплотнения удобно представить графически на плоскости частота-время (рис. 5.15). В системе с частотным уплотнением каждый сигнал передается по каналу все время и все сигналы смешиваются. Но каждый из них занимает свой конечный частотный интервал (не занятый каким-либо другим сигналом), как показано на рис. 5.15*а*. С другой стороны, в системе с временны́м уплотнением каждый сигнал занимает

свой временно́й интервал (не занятый другим сигналом). Но спектры всех сигналов лежат в одной и той же области (рис. 5.15*б*).

Выше было показано, что полоса частот, необходимая для передачи данного числа сигналов, одинакова при частотном и временно́м уплотнении (АИМ и АМ-ОБП требуют полосу nf_m, *гц*, АИМ/АМ и АМ-ДБП требуют полосу $2nf_m$, *гц*). Отсюда следует,

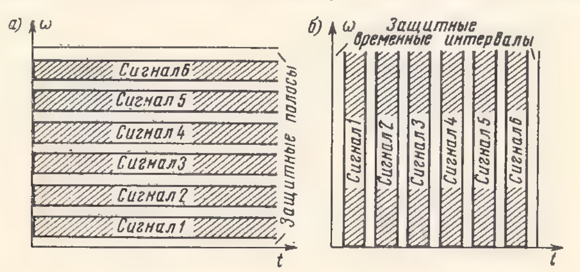


Рис. 5.15. Представление систем на частотно-временно́й плоскости:
а) с частотным уплотнением; *б)* с временны́м уплотнением

что в заданной полосе частот можно одновременно передать одинаковое число сигналов, используя частотное или временно́е уплотнение.

С практической точки зрения системы с временны́м уплотнением имеют преимущества перед системами с частотным уплотнением. Во-первых, системы с временны́м уплотнением проще систем с частотным уплотнением. В последних необходимо генерировать разные несущие для каждого канала. Кроме того, каждый канал занимает свою частотную полосу и, следовательно, требуются разные полосовые фильтры. С другой стороны, в системах с временны́м уплотнением требуется совершенно одинаковое оборудование в каждом канале, состоящее из сравнительно простых синхронных коммутаторов. Единственные фильтры, необходимые для детектирования, — это фильтры нижних частот, которые одинаковы во всех каналах. Эти устройства значительно проще модуляторов, демодуляторов, генераторов несущих колебаний и полосовых фильтров в системах с частотным уплотнением.

Второе преимущество систем с временны́м уплотнением заключается в их относительной нечувствительности к искажениям в канале, которые возникают в системах с частотным уплотнением вследствие нелинейностей в усилителях. Нелинейности в усилителях приводят к появлению гармоник передаваемых сигналов, а это создает помеху для других каналов (переходная помеха). Поэтому требования к нелинейности тракта передачи в системах с частотным уплотнением значительно более жесткие, чем при одноканальной передаче. С другой стороны, в системах с временны́м уплотнением сигналы разных каналов передаются не одновременно, а в отведенные для их передачи различные временны́е интерва-

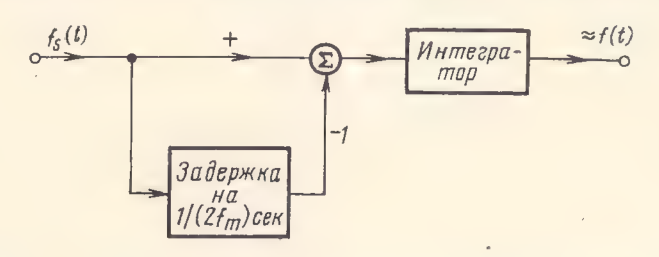
лы. Поэтому требования к нелинейности тракта передачи в многоканальной системе с временным уплотнением такие же, как в одноканальной системе. По этим причинам в таких применениях, как телефонная связь на большие расстояния, обычно используются системы с временным уплотнением.

ЗАДАЧИ

5.1. Перечисленные ниже сигналы не являются сигналами с ограниченным спектром. Однако приближенно их можно представить как сигналы с ограниченным спектром. Примите подходящий критерий для такой аппроксимации в каждом случае и определите минимальную частоту отсчетов:

$$\text{а)}\quad e^{-2|t|},\quad \text{б)}\quad e^{-2t}\cos 100t\,u_1(t),\quad \text{в)}\quad t\,e^{-t}u_1(t),\quad \text{г)}\quad G_{20}(t).$$

5.2. Если $f(t)$ — непрерывный сигнал, спектр которого ограничен частотой f_m, *гц*, а $f_s(t)$ — импульсный сигнал [выборка равномерная через $1/(2f_m)$, *сек*], то $f(t)$ можно восстановить из $f_s(t)$, пропустив его через фильтр нижних частот. На практике для этой цели используется цепь, структурная схема которой приведена на рисунке. Напряжение на выходе этой цепи приближенно соответствует сигналу $f(t)$. Изобразите форму колебаний в различных точках этой цепи для типичного импульсного сигнала $f_s(t)$.

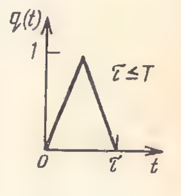


Какую передаточную функцию имеет эта цепь?

Указание. Определите импульсную реакцию цепи.

Представьте частотную характеристику этой цепи и сравните ее с характеристикой идеального фильтра нижних частот.

5.3. Сигнал $f(t)$ имеет спектр, ограниченный частотой f_m, *гц*. На рис. 5.2 показана естественная выборка сигнала $f(t)$ с использованием прямоугольных импульсов. Определите спектр импульсного сигнала, если естественная выборка выполняется с использованием импульсов произвольной формы $q(t)$, имеющих спектр $Q(\omega)$. Примите, что ширина импульса меньше интервала выборки. В частности, если $q(t)$ — треугольный импульс, показанный на рисунке, определите и изобразите выборочный сигнал и его спектр.

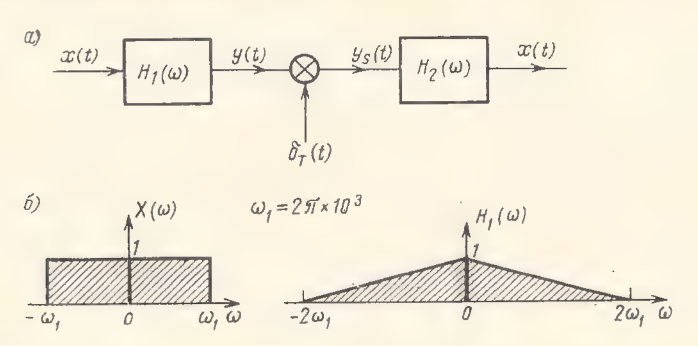


5.4. Сигнал $f(t)$ имеет спектр, ограниченный частотой f_m, *гц*. Выборка осуществляется импульсами треугольной формы, показанными на рисунке к задаче 5.3. Ширина импульса меньше чем интервал выборки T. Определите и изобразите спектр импульсного сигнала для случая мгновенной выборки.

5.5. Сигнал $x(t)$, имеющий спектр $X(\omega)$, показанный на рисунке (*б*), передается через цепь с передаточной функцией $H_1(\omega)$. Затем получается выборка с равномерной скоростью при помощи единичных импульсов. Каковы частоты выборки f_s, достаточные для точного восстановления сигнала $y(t)$?

Выберите одно из этих значений f_s; найдите и изобразите спектр импульсного сигнала $y_s(t)$.

Импульсный сигнал $y_s(t)$ затем пропускается через цепь с передаточной функцией $H_2(\omega)$. Какой должна быть функция $H_2(\omega)$ цепи, чтобы ее выходное напряжение в точности равнялось $x(t)$?



5.6. Проведенный в гл. 5 анализ ограничивался случаем, когда используемые при выборке импульсы имели меньшую ширину, чем интервал выборки T. Таким образом, все импульсы выборочного сигнала не перекрываются. Это ограничение, однако, не является необходимым. Покажите, что можно использовать перекрывающиеся импульсы (ширина которых больше, чем интервал выборки) при естественной выборке и восстановить исходный сигнал точно. Покажите, что процесс восстановления исходного сигнала из выборочного совпадает с процессом восстановления, используемым при естественной выборке, когда длительность импульсов меньше интервала выборки. Определите и дайте график выборочного сигнала и его спектра, если при выборке используются экспоненциальные импульсы $q(t)=e^{-t/T}$.

5.7. Повторите задачу 5.6 для случая мгновенной выборки.

5.8. Энергия импульса $q(t)$, используемого при мгновенной выборке, равна E_q. Определите среднюю мощность выборочного сигнала $x_s(t)$, выразив ее через E_q и $\overline{x^2(t)}$. Примите, что ширина импульсов меньше, чем интервал выборки T.

Указание. Используйте полученную в примере 2.2 ф-лу (2.31).

5.9. Постройте график функции $S_a(100\pi t)$. Ширина спектра этой функции равна 50 *гц*, а минимальная частота выборки — 100 *имп/сек*. Дайте график отсчетов этого сигнала, взятых с минимальной частотой выборки, причем выборка начинается при $t=0$.

Рассмотрите случаи, когда частота выборки равна 50 *имп/сек* и 25 *имп/сек*. Постройте график отсчетов при этих частотах выборки. Вы найдете, что отсчеты, полученные в этих трех случаях, одинаковы. Как это можно было бы объяснить?

Указание. Выборочный сигнал $f_s(t)=f(t)\delta_T(t)$, где $T=1/100$, $1/50$ и $1/25$ *сек*. Найдите спектр $f_s(t)$, используя теорему частотной свертки и покажите, что в каждом случае $f_s(t)$ содержит полную информацию о сигнале $f(t)$. Этот пример является исключением из общего правила.

5.10. Обобщите теорему равномерной выборки на случай сигналов, имеющих спектры, ограниченные полосой f_m, *гц*, но с центральной частотой, отличающейся от $\omega=0$. На положительных частотах спектр лежит от f_l до f_h, где $f_h-f_l=f_m$. Покажите, что минимальная частота выборки для таких сигналов должна быть $2f_h/n$, *имп/сек*, где f_h — верхняя частота спектра, а n — наибольшее целое число, меньшее f_h/f_m.

ГЛАВА

6

Шумы

6.1. Введение

В процессе передачи сигналы всегда подвергаются некоторым искажениям, которые называются шумами. В этой книге ограничимся рассмотрением случайных, т. е. непредсказуемых шумов. Такие шумы, как фон источника питания, самовозбуждение в системе с обратной связью и т. д., не являются случайными. Они могут быть устранены при правильном расчете аппаратуры и здесь рассматриваться не будут.

Источники шумов могут быть самыми различными. В зависимости от источника шумы делятся на искусственные, нерегулярные естественные и флуктуационные. Искусственные шумы возникают во всякого рода электрических устройствах: поврежденных контактах, электроустановках, системах зажигания, приборах с флуоресцентным свечением и др. Эти шумы всегда можно устранить, убрав источник шума. Нерегулярные естественные шумы возникают благодаря молниям, электрическим бурям в атмосфере, космическим излучениям и др. Флуктуационные шумы также являются естественными и возникают в физических системах благодаря случайным флуктуациям, таким, как тепловое движение свободных электронов в резисторах, случайная эмиссия электронов в вакуумных лампах, случайное возникновение, рекомбинация и диффузия носителей (дырок и электронов) в полупроводниках. В основном различают два вида флуктуационных шумов: дробовой и тепловой.

В настоящей главе шумы будут рассматриваться без подробного описания их статистических свойств.

6.2. Дробовой шум

ПРИРОДА ДРОБОВОГО ШУМА

Дробовой шум присутствует как в вакуумных лампах, так и в полупроводниковых приборах. В вакуумных лампах дробовой шум возникает вследствие случайного характера эмиссии электронов из катода. В полупроводниковых приборах этот эффект возникает

благодаря случайной диффузии неосновных носителей и случайному возникновению и рекомбинации пар дырка — электрон.

Природу дробового шума можно проиллюстрировать на примере электронной эмиссии с горячего катода плоско-параллельного диода, изображенного на рис. 6.1. При заданной температуре среднее число электронов, эмитируемых в секунду, постоянно. Однако процесс электронной эмиссии является случайным. Это означает, что если разделить временну́ю ось на большое число малых интервалов по $\Delta\tau$ *сек* каждый, количество электронов, эмитируемых в течение каждого такого интервала, окажется не постоянным, а случайным. В среднем же скорость эмиссии остается постоянной

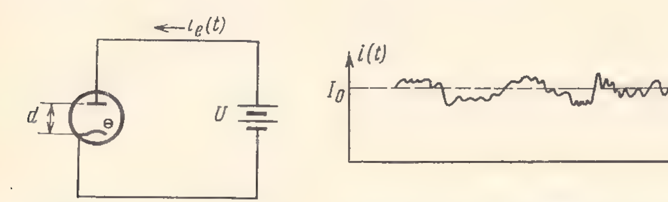


Рис. 6.1. Схема диода

Рис. 6.2. Осциллограмма тока диода при большой скорости развертки

при условии, что усреднение выполняется за достаточно большой промежуток времени. Таким образом, ток, образованный эмитируемыми электронами, не является постоянным и флуктуирует относительно среднего значения. Если наблюдать этот ток на экране осциллографа с медленной разверткой, он покажется постоянным. Однако, если этот ток наблюдать при быстрой развертке осциллографа, когда временная шкала сильно растянута, станет заметна случайная природа тока. Полный ток $i(t)$ можно рассматривать как сумму постоянного тока I_0 и шумового тока $i_n(t)$ с нулевым средним значением (рис. 6.2):

$$i(t)=I_0+i_n(t). \tag{6.1}$$

Природу флуктуаций тока $i(t)$ легче уяснить, если рассмотреть процесс наведения тока в аноде диода благодаря эмиссии электронов. Предположим, что из катода эмитируется один электрон (рис. 6.1). Этот электрон приобретает скорость при движении к аноду и наводит в анодной цепи ток $i_e(t)$. Если расстояние между катодом и анодом равно d единиц, эмитированный электрон находится под воздействием силы, направленной к аноду и равной qU/d, где q — заряд электрона, а U — приложенное напряжение. Электрон получает ускорение qU/md единиц, где m — масса электрона. Начальная скорость эмитируемого электрона обычно намного меньше конечной скорости, с которой он достигает анода. Допустим, что начальная скорость электрона равна нулю. Скорость $v(t)$ в произвольный момент времени t:

$$v(t)=\frac{qU}{md}t. \qquad (6.2)$$

Кинетическая энергия электрона в любой момент времени t равна $\frac{1}{2}mv^2$ или

$$Э=\frac{q^2U^2}{2md^2}t^2. \qquad (6.3)$$

Если движение этого электрона наводит на аноде заряд Q, то работа W, затрачиваемая на наведение заряда Q на аноде с потенциалом $U, W=QU$. Приравнивая эту работу кинетической энергии электрона, получим $Q=q^2Ut^2/2md^2$. Ток

$$i_e(t)=\frac{dQ}{dt}=\frac{q^2U}{md^2}t. \qquad (6.4)$$

Используя ф-лу (6.2), можно записать

$$i_e(t)=\frac{q}{d}v(t) \qquad (6.5)$$

Заметим, что наведенный ток пропорционален скорости электрона.

Время, необходимое электрону для достижения анода, называется временем пролета τ_a и легко определяется из ф-лы (6.2):

$$\left.\begin{aligned} d &= \frac{1}{2}\frac{qU}{md}\tau_a^2 \\ \tau_a &= \sqrt{\frac{2m}{qU}}\,d \end{aligned}\right\}. \qquad (6.6)$$

Подставив ф-лу (6.6) в (6.4), получим

$$i_e(t)=\begin{cases}\frac{2q}{\tau_a^2}t & (0>t<\tau_a) \\ 0 & (t>\tau_a).\end{cases} \qquad (6.7)$$

Очевидно, наведенный ток обращается в нуль, как только электрон достигает анода в момент $t=\tau_a$. Импульс тока, наведенный одиночным электроном, показан на рис. 6.3*а*. Общий ток состоит поэтому из большого числа таких треугольных импульсов, распределенных случайно во времени (рис. 6.3*б*). Сумма всех этих импульсов образует ток диода $i(t)$, показанный на рис. 6.2.

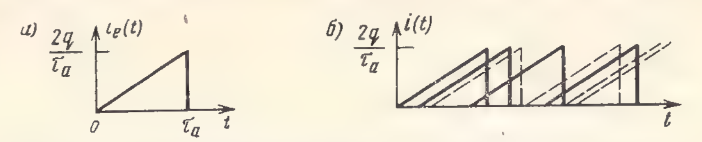


Рис. 6.3. К пояснению случайного характера тока диода: *а*) импульс тока, обусловленный одним электроном; *б*) импульсы тока, обусловленные множеством электронов

Заметим, что площадь под каждым импульсом равна q единиц. Следовательно, среднее значение анодного тока

$$I_0 = \overline{n} q, \tag{6.8}$$

где $\overline{n}$ — среднее число электронов, эмитируемых катодом за 1 *сек*.

СПЕКТР ПЛОТНОСТИ МОЩНОСТИ ДРОБОВОГО ШУМА

Обратим внимание на выражение тока дробового шума $i(t)$. Этот ток содержит две составляющие: постоянную составляющую I_0 и переменную составляющую $i_n(t)$. Составляющая $i_n(t)$, будучи случайной, не может быть записана как функция времени. Однако $i_n(t)$ представляет собой стационарный случайный сигнал, обладающий некоторым спектром плотности мощности. Так как в секунду на анод попадает $\overline{n}$ импульсов, можно ожидать, что спектр плотности мощности сигнала $i_n(t)$ равен спектру плотности энергии импульса $i_e(t)$, помноженному на $\overline{n}$ [1]. Таким образом, если

$$i_e(t) \leftrightarrow I_e(\omega), \tag{6.9}$$

то $S_i(\omega)$ — спектр плотности мощности тока $i_n(t)$:

$$S_i(\omega) = \overline{n} \, | I_e(\omega) |^2. \tag{6.10}$$

Функция $I_e(\omega)$ есть преобразование Фурье $i_e(t)$, и ее можно найти следующим образом:

$$i_e(t) = \frac{2q}{\tau_a^2} [t u_1(t) - \tau_a u_1(t - \tau_a) - (t - \tau_a) u_1(t - \tau_a)]. \tag{6.11}$$

Беря преобразование Лапласа от обеих частей равенства (6.11) и подставляя $i\omega$ вместо s, получаем

$$i_e(t) \leftrightarrow I_e(\omega) = \frac{2q}{-\omega^2 \tau_a^2} \left[1 - e^{-i\omega\tau_a} - i\omega\tau_a e^{-i\omega\tau_a} \right]. \tag{6.12}$$

Подставив ф-лу (6.12) в (6.10), получим

$$S_i(\omega) = \overline{n} \, | I_e(\omega) |^2 = \frac{4 I_0 q}{(\omega\tau_a)^4} [(\omega\tau_a)^2 + 2(1 - \cos\omega\tau_a - \omega\tau_a \sin\omega\tau_a)]. \tag{6.13}$$

Спектр плотности мощности $S_i(\omega)$ можно изобразить в функции ω. Как следует из ф-лы (6.13), более удобен график $S_i(\omega)$ в функции $\omega\tau_a$ (рис. 6.4). Заметим, что спектр плотности мощности оказывается почти равномерным для $\omega\tau_a < 0{,}5$.

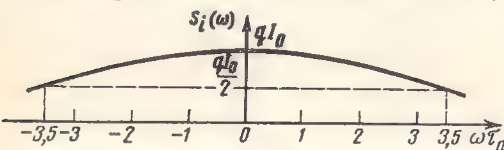


Рис. 6.4. Спектр плотности мощности дробового шума

---

[1] См., например, В. Б. Давенпорт, В. Л. Рут. Введение в теорию случайных сигналов и шумов. Гл. 7. Изд-во иностранной литературы, 1960. Lathi B. P. An Introduction to random signals and communication theory. International Textbook Co. Scranton, Pa., 1968.

Порядок величины τ_a можно вычислить на основании ф-лы (6.6). Отношение q/m для электрона равно $1{\cdot}76{\cdot}10^{11}$ *к/кг*. Следовательно,

$$\tau_a = 3{,}36 \cdot 10^{-6}\, d/\sqrt{U},\ \text{сек}.$$

Для диода с расстоянием между катодом и анодом $d=1$ *мм* (10^{-3} *м*) при $U=10$ *в* $\tau_a \approx 10^{-9}$ *сек*. Спектр плотности мощности шумовой составляющей тока в этом случае постоянен вплоть до $\omega \approx 0{,}5 \cdot 10^9 = 500 \cdot 10^6$ *рад/сек*. Это соответствует частоте около 80 *Мгц*. В общем случае спектр плотности мощности дробового шума можно считать постоянным до частоты 100 *Мгц*:

$$S_i(\omega) = qI_0. \qquad (6.14)$$

ДРОБОВОЙ ШУМ ДИОДА В РЕЖИМЕ ПРОСТРАНСТВЕННОГО ЗАРЯДА

Вольтамперную характеристику диода можно разбить на две области: область насыщения и область пространственного заряда (рис. 6.5). В области насыщения ток диода зависит от температуры катода. В этой области электрическое поле вполне достаточно для притяжения к аноду всех эмитируемых электронов. Поэтому увеличение напряженности электрического поля не может привести к увеличению тока I_0 (рис. 6.5). Средний ток может возрасти в этом случае только при увеличении температуры катода, что приведет к повышению эмиссии электронов с катода. На рис. 6.5 приведена вольтамперная характеристика типичного диода для трех значений температуры катода. При низком напряжении не все эмитируемые электроны достигают анода, некоторые из них остаются в пространстве между катодом и анодом, образуя так называемый пространственный заряд. Работа диода в этой области называется режимом пространственного заряда. В этом режиме ток диода I_0 можно увеличить за счет увеличения анодного напряжения.

Заметим, что в режиме насыщения дифференциальная проводимость $\partial I_0/\partial U$ диода почти равна нулю, т. е. диод обладает большим сопротивлением переменному току, тогда как в режиме пространственного заряда дифференциальная проводимость его не равна нулю. Спектр плотности мощности шума [ф-ла (6.14)] получен в предположении, что взаимодействие между электронами, летящими к аноду, отсутствует. Это предположение справедливо для режима насыщения, когда нет пространственного заряда. В этих условиях все электроны, эмитируемые катодом, достигают анода и анодный ток определяется только количеством электронов, эмитируемых в секунду, т. е. только температурой катода. Это имеет место при достаточно высоком анодном напряжении.

При низком напряжении в пространстве между анодом и катодом образуется электронное облако (пространственный заряд). Пространственный заряд снижает потенциал вблизи катода и от-

талкивает обратно некоторые эмитированные электроны. Несмотря на то, что электроны эмитируются катодом независимо друг от друга, после их вылета с катода между ними наблюдается сильное взаимодействие. Следовательно, возможность достижения электроном анода зависит от ранее эмитированных электронов. Из изложенного следует, что пространственный заряд ока-

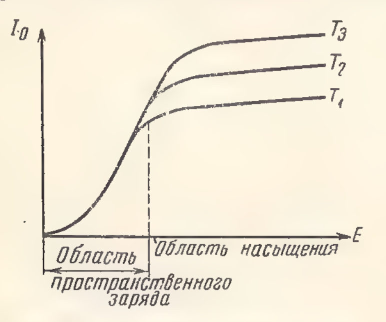


Рис. 6.5. Режимы работы диода

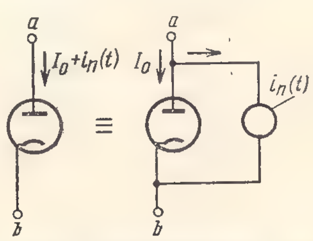


Рис. 6.6. Эквивалентная схема диода с источником шума

зывает сглаживающее действие на флуктуации тока. Например, при увеличении скорости эмиссии новые электроны увеличивают пространственный заряд, который теперь отталкивает бо́льшее число электронов к катоду. С другой стороны, при уменьшении скорости эмиссии пространственный заряд уменьшается и теперь бо́льшее число электронов достигает анода. Следовательно, пространственный заряд имеет тенденцию стабилизировать ток, сглаживая его флуктуации.

Можно показать[1]), что спектр плотности мощности шумового тока $i_n(t)$ в режиме пространственного заряда

$$S_i(\omega) = \alpha q I_0, \tag{6.15}$$

где

$$\alpha = 3\left(1 - \frac{\pi}{4}\right)\frac{2kT_c g_d}{qI_0} . \tag{6.16}$$

Следовательно,

$$S_i(\omega) = 1{,}288 kT_c g_d, \tag{6.17}$$

где T_c — абсолютная температура катода; $k = 1{,}38 \cdot 10^{-23}$ *дж/°К* — постоянная Больцмана; $g_d = \partial I_0 / \partial U$ — динамическая проводимость диода, зависящая от анодного напряжения[2]).

[1]) A. T. Rack. Effect of Space-Charge and Transit Time on the shot Noise in Diodes, Bell Syst. Tech. Journal, v. 17, pp. 592—619, 1938; R. T. Thompson. D. O. North. W. A. Harris. Fluctuations in Space-Charge Limited currents at Moderately High Frequences, RCA Rew. January 1940.

[2]) См. также В. Ф. Власов. Электронные и ионные приборы. Связьиздат, 1960. — *Прим. ред.*

Величина постоянной сглаживания α изменяется от 0,01 до 1. Итак,

$$S_i(\omega) = \begin{cases} qI_0 & \text{(в режиме насыщения)} \\ \alpha qI_0 & \text{(в режиме пространственного заряда)} \end{cases} \quad (6.18)$$

В общем случае диод можно представить как параллельное соединение бесшумового диода и источника шумового тока $i_n(t)$ (рис. 6.6), спектральная плотность мощности которого определяется соотношением (6.18).

ДРОБОВОЙ ШУМ В ТРИОДАХ И МНОГОЭЛЕКТРОДНЫХ ЛАМПАХ

Природа дробового шума в триодах, пентодах и других многоэлектродных вакуумных лампах в основном такая же, как и в диоде, работающем в режиме пространственного заряда. В триоде, например, напряжение U_g+U_a/μ играет ту же роль, что и напряжение между катодом и анодом в диоде. Плотность тока диода равна

$$J = CU^{3/2}, \quad (6.19)$$

где U — напряжение между катодом и анодом. Для триода, расстояние между сеткой и катодом которого равно расстоянию между анодом и катодом диода, эта формула записывается в виде[1])

$$J = C\sigma^{3/2}(U_g + U_a/\mu)^{3/2}, \quad (6.20)$$

где σ — коэффициент, учитывающий геометрические размеры лампы (обычно $0{,}5 \leqslant \sigma \leqslant 1$); U_g и U_a — соответственно сеточное и анодное напряжения относительно катода; μ — коэффициент усиления лампы.

Таким образом, диод с анодным напряжением $U=\sigma(U_g+U_a/\mu)$ будет иметь ту же плотность тока, что и триод с сеточным U_g и анодным U_a напряжениями. Следовательно, спектр плотности мощности шума триода на основании ф-лы (6.17)

$$S_i(\omega) = 1{,}288kT_c g_{экв},$$

где

$$g_{экв} = \frac{\partial I}{\partial[\sigma(U_g+U_a/\mu)]} = \frac{\partial I}{\partial U_g}\,\frac{\partial U_g}{\partial[\sigma(U_g+U_a/\mu)]} = \frac{g_m}{\sigma}$$

— взаимная проводимость эквивалентного диода;
g_m — динамическая проходная проводимость триода.

Следовательно,

$$S_i(\omega) = \frac{1{,}288kT_c g_m}{\sigma} = 2kT\,\frac{0{,}644T_c}{\sigma T}\,g_m. \quad (6.21)$$

---

[1]) K. R. Spangenberg. Vacuum Tube, Mc Graw-Hill, New Jork, 1948.

Обычно $T_c = 1000°K$, T — окружающая температура, равная 293°К и $\sigma = 0{,}088$. Поэтому

$$S_i(\omega) \approx 2kT(2{,}5g_m). \tag{6.22}$$

Таким образом, можно представить шумящий триод в виде параллельного соединения нешумящего триода и шумового источника тока $i_n(t)$, включенного между анодом и катодом (рис. 6.7б).

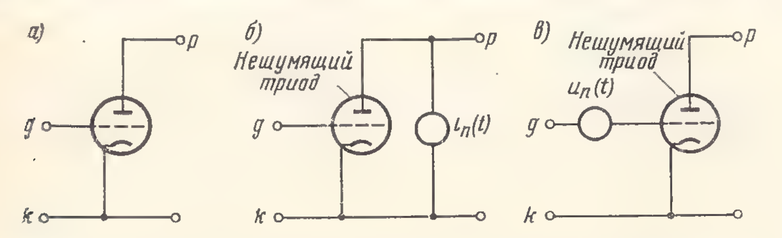


Рис. 6.7. Схемы представления шумящего триода:
а) триод; *б)* эквивалентная схема с параллельным источником шума; *в)* эквивалентная схема с последовательным источником шума в цепи сетки

Спектр плотности мощности тока $i_n(t)$ определяется по ф-ле (6.22). Если на сетке триода действует напряжение e_g, то анодный ток равен $g_m e_g$. Следовательно, можно считать, что ток дробового шума $i_n(t)$ возникает благодаря действию на сетке триода эквивалентного шумового напряжения

$$u_n(t) = i_n(t)/g_m. \tag{6.23}$$

Итак, шумящий триод можно представить как нешумящий, в цепь сетки которого включен шумовой источник напряжения (рис. 6.7*в*). Так как плотность мощности является функцией квадрата сигнала, то отношение спектров плотности мощности тока $i_n(t)$ и напряжения $u_n(t)$ равно g_m^2 (см. определение мощности и спектра плотности мощности). Следовательно,

$$S_u(\omega) = \frac{1}{g_m^2} S_i(\omega), \tag{6.24}$$

где $S_u(\omega)$ — спектр плотности мощности шума $u_n(t)$.

Из ф-л (6.22) и (6.24) получаем

$$S_u(\omega) = 2kT(2{,}5/g_m) = 2kTR_{\text{экв}}, \tag{6.25}$$

где

$$R_{\text{экв}} = 2{,}5/g_m. \tag{6.26}$$

Введение эквивалентного сопротивления $R_{\text{экв}}$ позволяет выразить дробовой шум в такой же форме, как и тепловой шум в резисторе (см. § 6.2).

Для пентодов и других многоэлектродных ламп справедливы эквивалентные схемы, приведенные для триодов. В многоэлектрод-

ных лампах возникают дополнительные шумы, вызванные случайным распределением катодного тока между разными электродами. Но и в этом случае дробовой шум и шум токораспределения можно заменить одним эквивалентным шумовым источником напряжения в цепи сетки со спектром плотности мощности $2kTR_{экв}$. Для пентодов

$$R_{экв} = 1 + \frac{7{,}7 I_s}{g_{mp}} R_{t\,экв}, \tag{6.27}$$

где g_{mp} — проходная проводимость пентода; I_s — экранный ток; $R_{t\,экв} = 2{,}5/g_{mt}$; g_{mt} — проходная проводимость пентода, включенного как триод.

В вакуумных лампах (а также и транзисторах) имеется еще один вид шума, так называемый фликкерный шум, имеющий спектр плотности мощности, пропорциональный $1/f$. Очевидно, такой шум будет преобладать на низких частотах (порядка нескольких килогерц). Этот шум возникает вследствие медленных изменений условий на поверхности катода и может быть уменьшен при соответствующей технологии изготовления поверхности катода вакуумной лампы и поверхностей переходов в транзисторах.

6.3. Тепловой шум

Тепловой шум возникает вследствие случайного движения свободных электронов в проводящей среде такой, как резистор. Каждый свободный электрон внутри резистора находится в движении благодаря своей тепловой энергии. Путь электронов является случайным и зигзагообразным вследствие столкновений с решетчатой структурой. Результирующим эффектом движения всех электронов является электрический ток, протекающий через резистор. Направление тока случайно и его среднее значение равно нулю. Можно показать[1]), что спектр плотности мощности тока свободных электронов

$$S_i(\omega) = \frac{2kTG\alpha^2}{\alpha^2 + \omega^2} = \tag{6.28}$$

$$= \frac{2kTG}{1 + (\omega/\alpha)^2}, \tag{6.29}$$

где k — постоянная Больцмана; T — температура окружающей среды; G — проводимость резистора (*сим*); α — среднее число столкновений электронов в секунду.

---

[1]) B. P. Lathi. An Infroduction to Random signals and Communication Theory. International Text book Co, Scranton, 1968, T. T. Freeman. Principles of Noise, Ch. 4. John Wiley and Sons., New York, 1958.

На рис. 6.8 приведен график спектра плотности мощности как функция ω/α. В пределах $\omega/\alpha<0{,}1$ спектр можно считать равномерным. Величина α имеет порядок 10^{14}, поэтому спектр оказывается равномерным до весьма высоких частот. Обычно спектр считают равномерным до частот порядка 10^{13} *гц*. Следовательно, для всех практически интересных случаев спектр плотности мощности теплового шума резистора можно определять по формуле

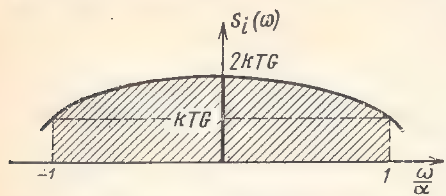


Рис. 6.8. Спектр плотности мощности теплового шума

$$S_i(\omega)=2kTG. \tag{6.30}$$

Поэтому вклад, который вносят тепловые шумы в любой ток, ограничивается только полосой пропускания цепи. Тепловой шум обычно рассматривается как шум, имеющий постоянный спектр плотности мощности. По этой причине тепловой шум называют также белым шумом.

Из изложенного следует, что резистор R можно представить как параллельное соединение нешумящей проводимости $G\,(G=1/R)$ и шумового источника тока i_n со спектром плотности мощности

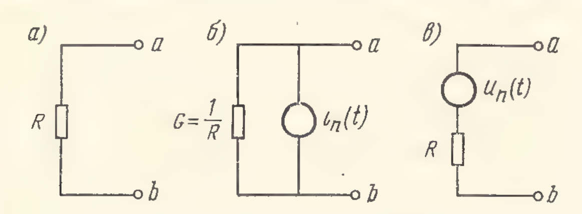


Рис. 6.9. Тепловые шумы в резисторе:
а) шумящий резистор; *б)* эквивалентная схема с параллельным источником шума; *в)* эквивалентная схема с последовательным источником шума

$2kTG$ (рис. 6.9*б*). Эквивалентная этому соединению схема замещения Тевенина (рис. 6.9*в*) состоит из последовательно включенных резистора R и шумового источника напряжения $u_n(t)=Ri_n(t)$. Так как спектр плотности мощности $S_u(\omega)$ является функцией квадрата сигнала, спектры плотности мощности $S_u(\omega)$ и $S_i(\omega)$ связаны соотношением

$$S_u(\omega)=R^2S_i(\omega)=R^2(2kTG)=2kTR. \tag{6.31}$$

Мощность спектральных составляющих шумового напряжения $u_n(t)$ в частотном диапазоне Δf, *гц*, вблизи центральной частоты ω (см. § 2.8)

$$\Delta P=2S_u(\omega)\,\Delta f=4kTR\,\Delta f. \tag{6.32}$$

Таким образом, произведение $4kTR$ есть мощность частотных составляющих $u_n(t)$ в единичной полосе (в герцах). Заметим, что по определению мощность сигнала есть среднее значение его квад-

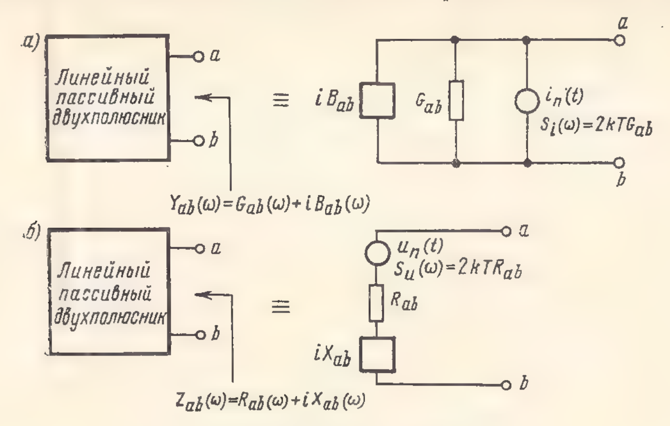


Рис. 6.10. Линейная пассивная цепь:
а) параллельная схема замещения; *б*) последовательная схема замещения

рата [см. ф-лу (2.21*б*)]. Если напряжение шума $u_n(t)$ проходит через узкополосный фильтр с полосой Δf, *гц*, то среднее значение квадрата выходного напряжения Δu_n

$$\overline{(\Delta u_n)^2} = 4kTR\,\Delta f. \tag{6.33}$$

Формулы (6.30) и (6.31) применимы для одиночных резисторов. Эти соотношения можно обобщить на любой линейный пассивный двухполюсник (например, содержащий элементы R, L и C). Такой двухполюсник может содержать несколько резисторов (рис. 6.10*а*), каждый из которых является источником теплового шума. Плотность мощности шумового напряжения в точках ab (рис. 6.10*а*), обусловленного всеми источниками, можно вычислить, обобщив соотношения (6.30) и (6.31).

Обобщенная теорема Найквиста, доказательство которой приводится ниже, устанавливает, что если проводимость цепи

$$Y_{ab} = G_{ab}(\omega) + \mathrm{i}\,B_{ab}(\omega),$$

то эту цепь можно представить как параллельное соединение нешумящей проводимости Y_{ab} и шумового источника тока со спектром плотности мощности $S_i(\omega)$ (рис. 6.10*а*):

$$S_i(\omega) = 2kTG_{ab}(\omega). \tag{6.34а}$$

Заметим, что эквивалентная проводимость $G_{ab}(\omega)$ в общем случае зависит от частоты, поэтому спектр плотности мощности эквивалентного шумового источника тока в общем случае является функцией частоты.

Эту теорему можно сформулировать иначе. Цепь с комплексным сопротивлением

$$Z_{ab}(\omega) = R_{ab}(\omega) + \mathrm{i}\, X_{ab}(\omega)$$

можно представить как последовательное соединение (рис. 6.10б) нешумящего комплексного сопротивления и шумового источника напряжения со спектром плотности мощности

$$S_u(\omega) = 2kTR_{ab}(\omega). \qquad (6.34б)$$

Важно помнить об ограничениях, накладываемых на эту теорему. Она применима к любому линейному двухполюснику, в котором все источники шума являются резисторами[1]). Таким образом, теорему можно применить к любой электрической цепи, состоящей из элементов R, L и C, но она неприменима к цепям с активными элементами, таким, как электронные лампы или транзисторы.

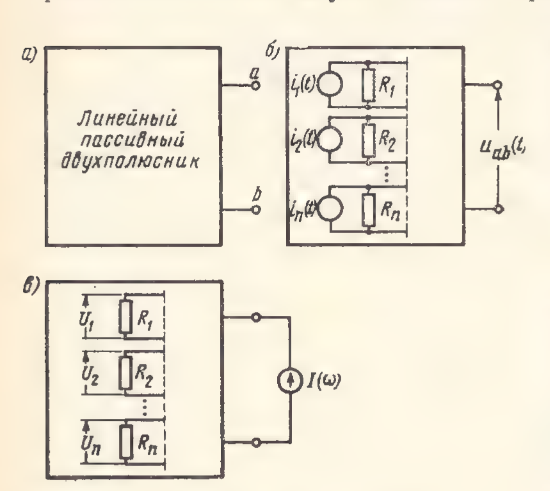


Рис. 6.11. К доказательству теоремы Найквиста:
а) линейный пассивный двухполюсник; б) двухполюсник с шумовыми источниками тока; в) двухполюсник с шумовыми источниками напряжения

Для доказательства теоремы Найквиста рассмотрим линейный пассивный двухполюсник (рис. 6.11а), содержащий n резисторов R_1, R_2, ... R_n. Каждый из этих резисторов представим нешумящим резистором, подключенным параллельно шумовому источнику тока, как это показано на рис. 6.11б.

Пусть $H_k(\omega)$ — передаточная функция для напряжения $u_{ab}(t)$ на клеммах ab относительно тока источника $i_k(t)$. Спектр плотности мощности источника тока

$$2kTG_k = 2kT/R_k.$$

Так как источники тока независимы, спектр плотности мощности выходного напряжения $S_o(\omega)$ на клеммах ab (см. § 6.4)

$$S_o(\omega) = \sum_{k=1}^{n} |H_k(\omega)|^2 \frac{2kT}{R_k} = 2kT \sum_{k=1}^{n} \frac{|H_k(\omega)|^2}{R_k}.$$

[1]) С помощью подходящего определения $R(\omega)$ эту теорему можно распространить на линейные диссипативные системы вообще, например на броуновское движение или флуктуации давления газа.

Покажем теперь, что если $Z_{ab}(\omega)$ — полное сопротивление в точках ab, то

$$R_{ab}(\omega) = \mathrm{Re}\,[Z_{ab}(\omega)] = \sum_{k=1}^{n} \frac{|H_k(\omega)|^2}{R_k}.$$

Из этого выражения следует, что

$$S_o(\omega) = 2\,k\,T\,R_{ab}(\omega).$$

Это и есть формулировка обобщенной теоремы Найквиста.

Докажем формулу для $R_{ab}(\omega)$. Рассмотрим ту же цепь при подаче на ее клеммы ab синусоидального тока с амплитудой I и частотой ω (рис. 6.11*в*). Обозначим этот ток через $I(\omega)$. Пусть синусоидальное напряжение на резисторе R_k, обусловленное током $I(\omega)$, будет $U_k(\omega)$. Так как система является линейным двухполюсником, то

$$U_k(\omega)/I(\omega) = H_k(\omega).$$

Поэтому мощность, выделяющаяся на резисторе R_k,

$$P_k = \frac{1}{2}\,\frac{|U_k(\omega)|^2}{R_k}.$$

Полная мощность P_t, выделяющаяся в цепи,

$$P_t = \frac{1}{2} \sum_{k=1}^{n} \frac{|U_k(\omega)|^2}{R_k}.$$

Используя равенство $H_k(\omega) = U_k(\omega)/I(\omega)$, получаем

$$P_t = \frac{1}{2} \sum_{k=1}^{n} \frac{|H_k(\omega)|^2}{R_k}\,|I(\omega)|^2 = \frac{1}{2}\,|I(\omega)|^2 \sum_{k=1}^{n} \frac{|H_k(\omega)|^2}{R_k}.$$

Полная мощность, рассеиваемая в цепи, равна полной мощности, доставляемой в цепь током $I(\omega)$. Но последняя равна $\frac{1}{2}|I(\omega)|^2 R_{ab}(\omega)$, где $R_{ab}(\omega) = \mathrm{Re}[Z_{ab}(\omega)]$. Очевидно,

$$\sum_{k=1}^{n} \frac{|H_k(\omega)|^2}{R_k} = R_{ab}(\omega) = \mathrm{Re}\,[Z_{ab}(\omega)].$$

Таким образом, формула для $R_{ab}(\omega)$ доказана и, следовательно, доказана и обобщенная теорема Найквиста.

Пример 6.1

Определим спектр плотности мощности напряжения теплового шума на клеммах ab цепи, показанной на рис. 6.12:

$$Y_{ab}(\omega) = 1 + \mathrm{i}\,\omega + \frac{1}{2 + 2\,\mathrm{i}\,\omega} = \frac{3 - 2\,\omega^2 + \mathrm{i}\,4\,\omega}{2 + \mathrm{i}\,2\,\omega}.$$

Следовательно,

$$Z_{ab}(\omega) = \frac{2 + \mathrm{i}\,2\,\omega}{3 - 2\,\omega^2 + \mathrm{i}\,4\,\omega}.$$

$$R_{ab}(\omega) = \mathrm{Re}\,[Z_{ab}(\omega)] = \frac{4\,\omega^2 + 6}{4\,\omega^4 + 4\,\omega^2 + 9}.$$

Спектр плотности мощности шумового напряжения $S_u(\omega)$ получим на основании эквивалентной схемы (рис. 6.12б)

$$S_u(\omega) = 2\,k\,TR_{ab}(\omega) = \frac{2\,k\,T\,(4\,\omega^2 + 6)}{4 + 4\,\omega^2 + 9}.$$

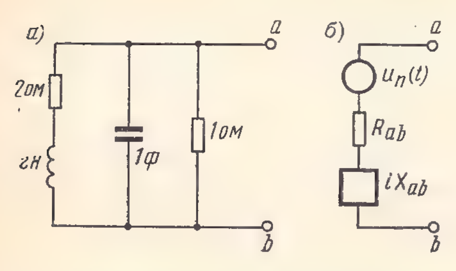


Рис. 6.12. К вычислению мощности шума в линейной цепи:
а) схема линейной цепи; *б*) последовательная эквивалентная схема

Если используется эквивалентная схема с шумовым источником тока (рис. 6.10*а*), то спектр плотности мощности шумового тока

$$S_i(\omega) = 2\,k\,T\,\mathrm{Re}\,[Y_{ab}(\omega)] = 2k\,T\,\frac{3 + 2\,\omega^2}{2 + 2\,\omega^2} = k\,T\,\frac{3 + 2\,\omega^2}{1 + \omega^2}.$$

6.4. Вычисление эффективного значения шума

ЦЕПЬ С ОДНИМ ИСТОЧНИКОМ ШУМА

Вычислим эффективное значение шума (т. е. квадратный корень из средней мощности шума) в произвольной цепи. Рассмотрим сначала цепь с одним источником шума. Пусть эта цепь состоит только из нешумящих элементов (рис. 6.13). К ее входу подключен источник шумового напряжения $u_n(t)$. Определим эффективное значение шумового напряжения на выходе $u_{no}(t)$. Пусть передаточная функция цепи, показанной на рис. 6.13, есть $H(\omega)$ и пусть $S_i(\omega)$ и $S_o(\omega)$ — спектры плотности мощности напряжений $u_{ni}(t)$ и $u_{no}(t)$ соответственно. Из ф-лы (2.37) следует, что

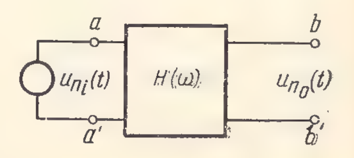


Рис. 6.13. Нешумящая цепь с источником шумового напряжения на входе

$$S_o(\omega) = S_i(\omega)\,|\,H(\omega)\,|^2.$$

Средняя мощность сигнала равна площади под кривой спектра плотности мощности, деленной на 2π,

$$\overline{u_{no}^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_o(\omega)\,d\,\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_i(\omega)\,|\,H(\omega)\,|^2\,d\,\omega. \qquad (6.35)$$

Заметим, что спектр плотности мощности — всегда четная функция ω, поэтому

$$\overline{u_{no}^2} = \frac{1}{\pi} \int_0^\infty S_i(\omega) \, | H(\omega) |^2 \, d\,\omega. \qquad (6.36а)$$

Эффективное значение выходного шума

$$\sqrt{\overline{u_{no}^2}} = \left[\frac{1}{\pi} \int_0^\infty S_i(\omega) \, | H(\omega) |^2 \, d\,\omega \right]^{1/2}. \qquad (6.36б)$$

Пример 6.2.

В качестве примера рассмотрим *RC*-цепь, изображенную на рис. 6.14. Вычислим эффективное напряжение шума на клеммах *aa'* конденсатора *C*.

Заменим резистор *R* нешумящим резистором и последовательно включенным с ним шумовым источником напряжения $u_{ni}(t)$ с плотностью мощности $2kTR$. Передаточная функция

$$H(\omega) = \frac{1/(\mathrm{i}\,\omega C)}{R + 1/(\mathrm{i}\,\omega RC)} = \frac{1}{\mathrm{i}\,\omega RC + 1}.$$

Рис. 6.14. К вычислению эффективного напряжения шума в *RC*-цепи:

а) схема *RC*-цепи; *б)* эквивалентная схема

Спектр плотности мощности выходного напряжения $u_{no}(t)$

$$S_o(\omega) = S_i(\omega) \, |H(\omega)|^2 = \frac{2\,kTR}{1 + \omega^2 C^2 R^2}.$$

Отсюда эффективное напряжение шума согласно ф-ле (6.36б)

$$\sqrt{\overline{u_{no}^2}} = \left[\frac{1}{\pi} \int_0^\infty \frac{2\,kTR}{1 + \omega^2 C^2 R^2} \, d\,\omega \right]^{1/2} = \left[\frac{2\,kTR}{\pi\, CR} \operatorname{arc\,tg}(\omega CR) \Big|_0^\infty \right]^{1/2} = \sqrt{\frac{kT}{C}}.$$

ЦЕПЬ С НЕСКОЛЬКИМИ ИСТОЧНИКАМИ ШУМА

В общем случае электрическая цепь может содержать не один, а большое число источников шума, таких, как резисторы и электронные лампы, причем все они являются независимыми источниками случайных сигналов. Покажем, что для множества независимых источников случайных сигналов при определении средней мощности и спектра плотности мощности применим принцип суперпозиции. Это означает, что в случае нескольких источников шума средняя мощность отклика системы равна сумме средних мощностей откликов, вычисленных в предположении, что действует один из рассматриваемых источников шума. Подобным образом,

спектр плотности мощности отклика равен сумме спектров плотности мощности, полученных в предположении, что действует один из рассматриваемых источников.

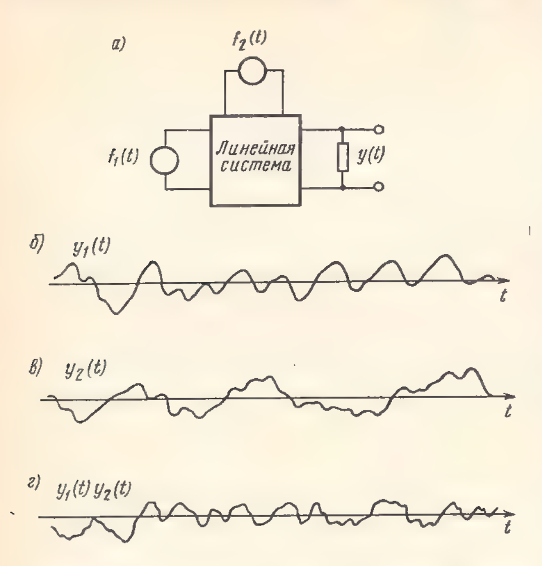


Рис. 6.15. Линейная цепь с двумя источниками шума:

а) схема; б) отклик системы на воздействие шума первого источника; в) отклик системы на воздействие шума второго источника; г) результат перемножения откликов

Рассмотрим систему (рис. 6.15*а*) с двумя источниками сигналов $f_1(t)$ и $f_2(t)$, подключенными к двум различным входам, и откликом $y(t)$. В силу линейности системы $y(t) = y_1(t) + y_2(t)$, где $y_1(t)$ — отклик, вызванный действием только $f_1(t)$, а $y_2(t)$ — отклик, вызванный только $f_2(t)$. Покажем, что если $f_1(t)$ и $f_2(t)$ являются независимыми источниками шумов с нулевыми средними значениями[1]), то

$$\overline{y^2(t)} = \overline{y_1^2(t)} + \overline{y_2^2(t)}, \quad (6.37а)$$

$$S_y(\omega) = S_{y1}(\omega) + S_{y2}(\omega). \quad (6.37б)$$

Это можно доказать следующим образом: $y(t) = y_1(t) + y_2(t)$, а потому

$$\overline{y^2(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} [y_1(t) + y_2(t)]^2\, dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} y_1^2(t)\, dt +$$

$$+ \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} y_2^2(t)\, dt + \lim_{T\to\infty} \frac{2}{T} \int_{-T/2}^{T/2} y_1(t)\, y_2(t)\, dt =$$

$$= \overline{y_1^2(t)} + \overline{y_2^2(t)} + 2\lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} y_1(t)\, y_2(t)\, dt. \quad (6.38)$$

Интеграл в правой части полученного выражения представляет собой среднее значение произведения $y_1(t)\,y_2(t)$. Если $y_1(t)$ и $y_2(t)$ — независимые случайные сигналы с нулевыми средними значениями (рис. 6.15*б* и *в*), то и $y_1(t)$ и $y_2(t)$ с равной вероятно-

[1]) Достаточно, чтобы один из сигналов имел нулевое среднее значение.

стью принимают как положительные, так и отрицательные значения. Очевидно, их произведение также с равным успехом принимает как отрицательные, так и положительные значения (рис. 6.15г), поэтому среднее значение произведения $y_1(t)y_2(t)$ должно быть равным нулю [1]). В устойчивой системе, если входная величина имеет нулевое среднее, то и выходная также имеет нулевое среднее значение [2]). Поэтому, если $x_1(t)$ и $x_2(t)$ имеют нулевые средние, то $y_1(t)$ и $y_2(t)$ также имеют нулевые средние значения. При этом соотношение (6.38) принимает вид

$$\overline{y^2}(t) = \overline{y_1^2}(t) + \overline{y_2^2}(t).$$

Среднее значение квадрата сигнала равно площади под кривой его спектра плотности мощности, деленной на 2π. Поэтому ф-лу (6.38) можно представить как

$$\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_y(\omega)\,d\omega = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_{y\,1}(\omega)\,d\omega + \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_{y\,2}(\omega)\,d\omega =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} [S_{y\,1}(\omega) + S_{y\,2}(\omega)]\,d\omega.$$

Последнее равенство можно записать в виде

$$S_y(\omega) = S_{y\,1}(\omega) + S_{y\,2}(\omega). \qquad (6.39)$$

Следовательно, спектр плотности мощности подчиняется принципу суперпозиции, который иллюстрируется соотношениями (6.37) для средних мощностей и спектров плотности мощности двух независимых случайных сигналов с нулевыми средними значениями [3]).

---

[1]) Легко видеть, что этот вывод справедлив, если только $y_1(t)$ либо $y_2(t)$ имеет нулевое среднее значение.

[2]) Если система устойчива, то импульсная реакция $h(t)$ затухает во времени и среднее значение $h(t)$ равно нулю, что означает отсутствие у $h(t)$ постоянной составляющей. Следовательно, $H(\omega)$ не содержит единичного импульса в начале координат и значение $|H(\omega)|^2$ конечно при $\omega=0$. Если $x(t)$ — входной, а $y(t)$ — выходной сигналы, то

$$S_y(\omega) = |H(\omega)|^2 S_x(\omega).$$

Если $x(t)$ имеет нулевое среднее, $S_x(\omega)$ не содержит единичного импульса при $\omega=0$. Следовательно, $|H(\omega)|^2 S_x(\omega)$ конечно при $\omega=0$ и не содержит единичного импульса при $\omega=0$. Очевидно, $y(t)$ не имеет постоянной составляющей, т. е. среднее значение отклика равно нулю.

[3]) Можно показать, что принцип суперпозиции спектров плотности мощности применим не только к некоррелированным сигналам, но и к независимым. Независимые источники являются некоррелированными, но не наоборот. Это невозможно доказать без привлечения некоторых положений теории случайных процессов.

Таким образом, при нескольких независимых случайных источниках в системе спектральную плотность мощности можно получить, предполагая на время наличие одного источника и вычисляя спектр плотности мощности отклика. Результирующая плотность мощности, вызванная всеми источниками, равна сумме всех отдельных плотностей мощности.

Заметим, что принцип суперпозиции спектров плотности мощности применим к независимым случайным источникам[1]). В этом случае интеграл в ф-ле (6.38) равен нулю и считается, что шумы отдельных источников суммируются некогерентно. Если же между источниками имеется статистическая связь, интеграл в ф-ле (6.38) не равен нулю и соотношение (6.37) несправедливо.

Спектр плотности мощности на выходе линейного RLC-двухполюсника, содержащего несколько резисторов, можно вычислить с помощью обобщенной теоремы Найквиста или принципа суперпозиции спектров плотности мощности. В общем случае, однако, теорема Найквиста более универсальна. Читатель может это проверить, вычислив с помощью принципа суперпозиции спектр плотности мощности напряжения на клеммах ab цепи, приведенной на рис. 6.12*а*.

Пример 6.3.

Вычислим эффективное значение шумового напряжения на выходных клеммах усилителя, показанного на рис. 6.16*а*. Чтобы решить эту задачу, заменим усилитель эквивалентной схемой, состоящей из нешумящих элементов и эквивалентных источников шума. Триод можно представить источником шума на выходе (рис. 6.7*б*) или источником шума на входе (рис. 6.7*в*). Второй вариант более удобен, чем первый, поэтому во всех дальнейших вычислениях будем

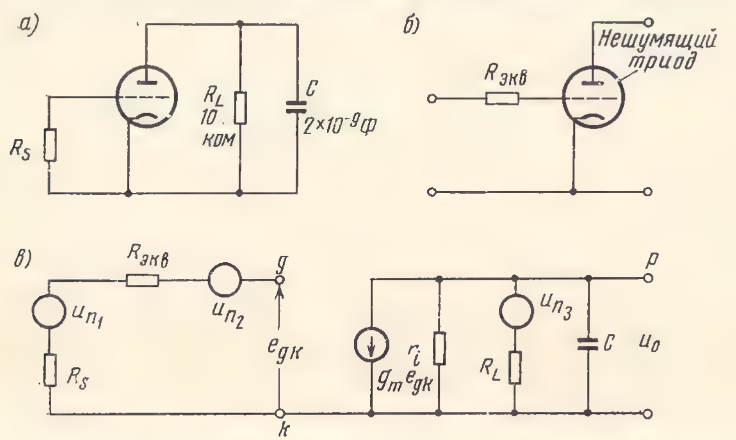


Рис. 6.16. К вычислению шумов в триодном усилителе: *а)* схема усилителя; *б)* схема с эквивалентным шумящим сопротивлением; *в)* полная эквивалентная схема

[1]) Точнее к некоррелированным случайным источникам.

заменять электронную лампу эквивалентной схемой с шумовым источником во входной цепи. Заметим, что эквивалентное напряжение шума на сетке лампы имеет спектр плотности мощности $2kTR_{экв}$, где $R_{экв}=2{,}5/g_m$. Такой источник можно заменить эквивалентным резистором $R_{экв}$ в сеточной цепи лампы, как это показано на рис. 6.16*б*. Плотность мощности теплового шума, генерируемого этим резистором, равна $2kTR_{экв}$ и равна, в свою очередь, эквивалентной плотности мощности дробового шума на сетке. Такая замена вполне удовлетворительна до частот порядка 100 *Мгц*, когда можно еще полагать плотность мощности дробового шума постоянной. Заметим, что включение эквивалентного резистора в цепь сетки не влияет на режим работы триода, так как в этой части схемы ток (постоянный) не протекает.

В данном усилителе присутствует шум от трех источников: тепловой шум резистора R_s; дробовой шум лампы, замененный эквивалентным тепловым шумом резистора $R_{экв}$ в сетке, и тепловой шум нагрузочного резистора R_L. Пусть $S_1(\omega)$, $S_2(\omega)$ и $S_3(\omega)$ — спектры плотности мощности шумовых напряжений этих трех источников. Очевидно,

$$S_1(\omega) = 2kTR_s = 2kT\cdot 10^3,$$

$$S_2(\omega) = 2kTR_{экв} = 2kT\frac{2{,}5}{g_m} = 2kT\cdot 10^3,$$

$$S_3(\omega) = 2kTR_L = 2kT\cdot 10^4.$$

На рис. 6.16*в* приведена полная эквивалентная схема усилителя, где триод представлен эквивалентным источником шумового тока. Три источника шумового напряжения обозначены соответственно u_{n1}, u_{n2} и u_{n3}. Выходное напряжение равно u_o. Если через $H_1(\omega)$, $H_2(\omega)$ и $H_3(\omega)$ обозначить передаточные функции, связывающие напряжение u_o с u_{n1}, u_{n2} и u_{n3} соответственно, то можно показать, что

$$H_1(\omega) = H_2(\omega) = \frac{-g_m}{1/r_i + 1/R_L + i\omega C} = \frac{-g_m r_i R_L}{R_L + r_i + i\omega r_i R_L C}$$

и

$$H_3(\omega) = \frac{r_i/(1 + i\omega r_i C)}{R_L + r_i/(1 + i\omega r_i C)} = \frac{r_i}{R_L + r_i + i\omega r_i R_L C}.$$

Спектр плотности мощности выходного напряжения

$$S_o(\omega) = |H_1(\omega)|^2 S_1(\omega) + |H_2(\omega)|^2 S_2(\omega) + |H_3(\omega)|^2 S_3(\omega) =$$

$$= \frac{g_m^2 r_i^2 R_L^2}{(R_L + r_i)^2 + \omega^2 r_i^2 R_L^2 C^2}(4kT\cdot 10^3) + \frac{r_i^2}{(R_L + r_i)^2 + \omega^2 r_i^2 R_L^2 C^2}(2kT\cdot 10^4) =$$

$$= \frac{2kT r_i^2\cdot 10^3}{(R_L + r_i)^2 + \omega^2 r_i^2 R_L^2 C^2}\left(2g_m^2 R_L^2 + 10\right).$$

Заметим, что относительная составляющая выходного шума, вызванного R_s и $R_{экв}$, равна $2g_m^2 R_L^2$ и составляет 1250, а составляющая, обусловленная R_L, равна только 10. Такое соотношение справедливо в большинстве практических случаев. Шумом резистора нагрузки можно пренебречь, так как он мал по сравнению с шумом лампы и входного резистора.

Подставив значения R_L, r_i, C, k и T (293°K) в ф-лу (6.39), получим

$$S_o(\omega) = \frac{2{,}52\cdot 10^{-5}}{10^{10} + \omega^2}. \tag{6.40}$$

Следовательно,

$$\overline{u_o^2}=\frac{1}{\pi}\int\limits_0^\infty S_o(\omega)\,d\omega=\frac{2{,}52\cdot 10^{-5}}{\pi}\int\limits_0^\infty \frac{d\omega}{10^{10}+\omega^2}=1{,}26\cdot 10^{-10}.$$

Эффективное значение напряжения шума на выходе

$$\sqrt{\overline{u_o^2}}=11{,}2\cdot 10^{-6}=11{,}2 \text{ мкв}.$$

6.5. Эквивалентная шумовая полоса

Иногда полезно знать эквивалентную шумовую полосу электрической цепи. Среднее значение квадрата выходного сигнала в системе с передаточной функцией $H(\omega)$ и входным спектром плотности мощности $S_i(\omega)$ определяется ф-лой (6.35). Большинство рассматриваемых цепей имеют достаточно узкие полосы пропускания, что позволяет считать спектральную плотность мощности теплового и дробового шума $S_i(\omega)$ постоянной в пределах полосы пропускания и равной, например, K. Таким образом,

$$\overline{u_o^2}=\frac{K}{2\pi}\int\limits_{-\infty}^{\infty}|H(\omega)|^2\,d\omega=\frac{K}{\pi}\int\limits_0^\infty |H(\omega)|^2\,d\omega. \tag{6.41}$$

Для заданной цепи интеграл в ф-ле (6.41) — величина постоянная. Эквивалентную шумовую полосу W_0 на некоторой частоте ω_0 определим следующим образом:

$$W_0=\frac{1}{|H(\omega_0)|^2}\int\limits_0^\infty |H(\omega)|^2\,d\omega. \tag{6.42}$$

Следовательно,

$$\overline{u_o^2}=K\,|H(\omega_0)|^2\,W_0. \tag{6.43}$$

Теперь ясен смысл понятия эквивалентной шумовой полосы. Среднее значение квадрата шумового сигнала на выходе данной системы совпадает с аналогичным параметром шумового сигнала идеальной системы с ограниченной полосой W_0 и постоянным коэффициентом передачи $H(\omega_0)$. Это иллюстрируется рис. 6.17.

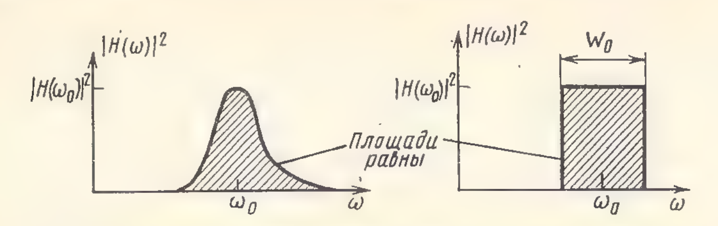


Рис. 6.17. К определению эквивалентной шумовой полосы цепи

6.6. Коэффициент шума усилителя

Из изложенного выше следует, что сигналы всегда искажаются шумами. Кроме того, при прохождении через любую систему к сигналу добавляется шум. Чистоту сигнала (или относительный уровень сигнала по сравнению с шумом) хорошо характеризует отношение мощности сигнала к мощности шума, которое называется отношением сигнал/шум.

При усилении сигнала усилителем к его первоначальным шумам добавляются шумы, источником которых является усилитель. Это вызывает изменение выходного отношения сигнал/шум по сравнению со входным. Отношение входного отношения сигнал/шум к выходному характеризует шумовые свойства усилителя. При прохождении через усилитель входные шумы усиливаются и выделяются на нагрузке. Дополнительный шум возникает в самом усилителе и также выделяется на нагрузке. Следовательно, шум на выходе создается источником входного шума и самим усилителем.

Определим коэффициент шума F усилителя как отношение полной плотности мощности шума в нагрузке (на выходе) к плотности мощности шума в нагрузке (на выходе), обусловленного только входными шумами. Если $S_{nto}(\omega)$ и $S_{nso}(\omega)$ представляют спектры плотности мощности полного шума на выходе и выходного шума, обусловленного только входным источником шума соответственно, то по определению

$$F = \frac{S_{nto}(\omega)}{S_{nso}(\omega)}. \qquad (6.44)$$

Если $S_{nao}(\omega)$ — спектр плотности мощности шума в нагрузке, обусловленный самим усилителем, то

$$S_{nto}(\omega) = S_{nso}(\omega) + S_{nao}(\omega),$$

$$F = \frac{S_{nso}(\omega) + S_{nao}(\omega)}{S_{nso}(\omega)} = 1 + \frac{S_{nao}(\omega)}{S_{nso}(\omega)}. \qquad (6.45)$$

Заметим, что сопротивление нагрузки также вносит свой вклад в выходные шумы и этот шум включается в S_{nao}. Однако величина этого шума обычно очень мала по сравнению с шумом, создаваемым активным прибором (см. пример 6.3) и источником шума на входе усилителя, и поэтому шумом нагрузки можно пренебречь. Очевидно, коэффициент шума усилителя является сравнительной мерой «шумности» усилителя по сравнению с «шумностью» источника.

Можно и в другом виде выразить коэффициент шума. Пусть $S_{si}(\omega)$ и $S_{so}(\omega)$ — спектры плотности мощности сигнала на входе и выходе усилителя соответственно. Тогда

$$S_{so}(\omega) = S_{si}(\omega)\,|H(\omega)|^2. \qquad (6.46)$$

Если $S_{nsi}(\omega)$ — спектр плотности мощности шумового источника (входного шума), то

$$S_{nso}(\omega) = S_{nsi}(\omega)\,|H(\omega)|^2$$

и по определению (6.44)

$$F = \frac{S_{nto}(\omega)}{S_{nsi}(\omega)\,|H(\omega)|^2}.$$

Подстановка ф-лы (6.46) в это соотношение дает

$$F = \frac{S_{si}(\omega)/S_{nsi}(\omega)}{S_{so}(\omega)/S_{nto}(\omega)}. \qquad (6.47)$$

Числитель в полученной формуле представляет собой отношение плотностей мощности сигнала и шума на входе, а знаменатель — отношение плотностей мощности сигнала и шума на выходе. Следовательно, коэффициент шума измеряет ухудшение отношения плотностей мощности сигнала и шума в процессе усиления. Согласно этому определению очевидно, что коэффициент шума всегда больше единицы и, следовательно, отношение плотностей мощности сигнала и шума всегда ухудшается в процессе усиления. Это ухудшение происходит, конечно, за счет шумов усилителя.

Коэффициент шума F, определенный соотношением (6.44) или (6.47), зависит от частоты и поэтому иногда называется спектральным коэффициентом шума. С другой стороны, можно определить средний коэффициент шума $\overline{F}$ как отношение полной мощности шумов на выходе к мощности шума, вносимого только входным источником.

Заметим, что это определение является просто некоторым обобщением определения (6.44). В этом определении фигурирует общая мощность шума на всех частотах, поэтому оно соответствует интегральному или среднему коэффициенту шума. Таким образом, по определению

$$\overline{F} = \frac{\frac{1}{\pi}\int\limits_0^\infty S_{nto}(\omega)\,d\omega}{\frac{1}{\pi}\int\limits_0^\infty S_{nso}(\omega)\,d\omega} = 1 + \frac{\int\limits_0^\infty S_{nao}(\omega)\,d\omega}{\int\limits_0^\infty S_{nso}(\omega)\,d\omega} = 1 + \frac{N_{ao}}{N_{so}}, \qquad (6.48)$$

где N_{ao} — мощность шума на выходе, обусловленного только усилителем; N_{so} — мощность шума на выходе, обусловленного только источником шума.

Из определения коэффициента шума (спектрального или среднего) следует, что коэффициент шума всегда больше единицы и равен единице в случае идеального (нешумящего) усилителя. Близость коэффициента шума усилителя к единице является показателем качества усилителя (с точки зрения шумовых свойств). Необходимо, однако, подчеркнуть, что коэффициент шума характери-

зует не абсолютное, а относительное качество усилителя. Он показывает шумовые свойства усилителя по сравнению с «шумностью» источника.

Из определения следует, что коэффициент шума усилителя можно сделать весьма близким к единице, увеличивая шум внешнего источника. Конечно, такое решение не улучшает качество усилителя, так как в этом случае источник настолько сильно шумит по сравнению с усилителем, что последний можно считать почти нешумящим. При этом сильно ухудшается общее отношение сигнал/шум на выходе и выходной сигнал оказывается сильно искаженным шумами. Поэтому для улучшения коэффициента шума не нужно увеличивать шумы источника (или уменьшать отношение сигнал/шум на входе). Во многих случаях задача решается посредством включения повышающего трансформатора, который усиливает как входной шум, так и сигнал. Увеличение шума на входе приводит к тому, что усилитель оказывается как бы менее шумящим без ухудшения входного отношения сигнал/шум. Поэтому коэффициент шума уменьшается и выходное отношение сигнал/шум действительно улучшается (см. пример 6.5).

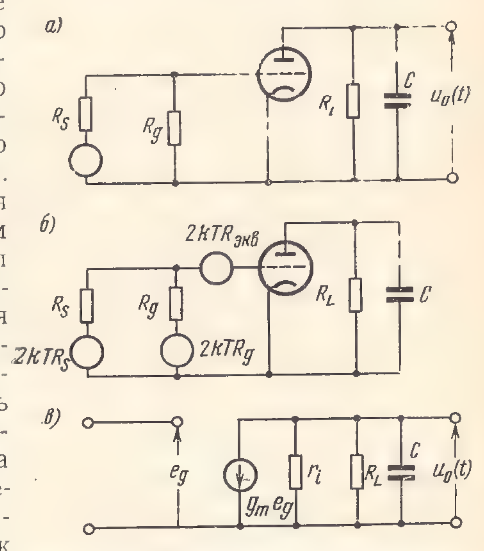


Рис. 6.18. К определению коэффициента шума триодного усилителя:
а) схема усилителя; *б)* схема с выделенными источниками шума; *в)* эквивалентная схема

Пример 6.4.

Определим коэффициент шума триодного усилителя, показанного на рис. 6.18*а*, и вычислим эффективное значение напряжения шума на выходе усилителя. Дано: внутреннее сопротивление лампы $r_i = R_L = 10$ *ком*, $g_m = 2{,}5 \cdot 10^{-3}$ *сим*, $R_g = 100$ *ком*, $R_s = 1$ *ком* и $C = 2 \cdot 10^{-9}$ *ф*.

Шум сопротивления нагрузки обычно пренебрежимо мал по сравнению с другими составляющими шума, поэтому R_L можно полагать нешумящим (см. пример 6.3).

Сопротивления источника R_s и цепи сетки R_g представим нешумящими резисторами, включенными последовательно с эквивалентными шумовыми источниками напряжения, а триод представим нешумящим триодом и эквивалентным шумовым источником напряжения в сетке (рис. 6.18*б*).

Спектры плотности мощности, обусловленные R_s, R_g и $R_{экв}$, равны соответственно $2kTR_s$, $2kTR_g$ и $2kTR_{экв}$, как показано на рис. 6.18*б*. Пусть $H(\omega)$ — передаточная функция, связывающая выходное напряжение с сеточным. Отсюда следует, что передаточная функция для эквивалентного шума, обусловленно-

го сопротивлением $R_{экв}$, равна $H(\omega)$, а передаточная функция для шума, создаваемого сопротивлением R_g, есть $\frac{R_s}{R_s+R_g}H(\omega)$.

Подобным образом передаточная функция для напряжения шума, создаваемого сопротивлением R_s, равна $\frac{R_g}{R_s+R_g}H(\omega)$

Следовательно, спектр плотности мощности шума на выходе, обусловленного $R_{экв}$, равен

$$S_{R_{экв}\,o}(\omega) = 2\,kTR_{экв}\,|H(\omega)|^2. \tag{6.49а}$$

Спектры плотности мощности шума на выходе, создаваемого сопротивлениями R_g и R_s, равны соответственно

$$S_{R_g\,o}(\omega) = 2\,kTR_g\left(\frac{R_s}{R_s+R_g}\right)^2|H(\omega)|^2 \tag{6.49б}$$

и

$$S_{R_s\,o}(\omega) = 2kT\,R_s\left(\frac{R_g}{R_s+R_g}\right)^2|H(\omega)|^2. \tag{6.49в}$$

Заметим, что сеточное сопротивление R_g является элементом усилителя, и поэтому

$$S_{nao}(\omega) = 2\,kTR_{экв}\,|H(\omega)|^2 + 2\,kTR_g\left(\frac{R_s}{R_s+R_g}\right)^2|H(\omega)|^2 \tag{6.50а}$$

и

$$S_{nso}(\omega) = 2\,kTR_s\left(\frac{R_g}{R_s+R_g}\right)^2|H(\omega)|^2. \tag{6.50б}$$

Следовательно, коэффициент шума

$$F = 1 + \frac{S_{nao}}{S_{nso}} = 1 + \frac{R_{экв} + R_g\left(\frac{R_s}{R_s+R_g}\right)^2}{R_s\left(\frac{R_g}{R_s+R_g}\right)^2} =$$

$$= 1 + \frac{R_{экв}(R_s+R_g)^2 + R_g R_s^2}{R_s R_g^2}, \tag{6.51}$$

где $R_{экв} = 2{,}5/g_m = 2{,}5/2{,}5\cdot10^{-3} = 1000$ *ом* и $R_s = 1000$ *ом* и $R_g = 10^5$ *ом*.

Подстановка этих значений сопротивлений в ф-лу (6.51) дает $F=2{,}03$. Для того чтобы найти эффективное значение выходного напряжения шума, необходимо определить передаточную функцию $H(\omega)$. На рис. 6.18*в* приведена эквивалентная схема триода, с помощью которой можно определить передаточную функцию, связывающую выходное напряжение с сеточным:

$$H(\omega) = -g_m Z,$$

где

$$\frac{1}{Z} = \frac{1}{r_i} + \frac{1}{R_L} + i\,\omega\,C = 10^{-4} + 10^{-4} + (2\cdot10^{-9})(i\,\omega) = 2\cdot10^{-9}\,(10^5 + i\,\omega)$$

$$H(\omega) = \frac{2{,}5\cdot10^{-3}}{(2\cdot10^{-9})(10^5+i\,\omega)} = \frac{1{,}25\cdot10^6}{10^5+i\,\omega},$$

$$|H(\omega)|^2 = \frac{1{,}56\cdot10^{12}}{10^{10}+\omega^2}. \tag{6.52}$$

Средняя мощность напряжения шума на выходе равна

$$\frac{1}{\pi}\int_0^\infty S_{nto}(\omega)\,d\omega,$$

где

$$S_{nto}(\omega) = S_{nao}(\omega) + S_{nso}(\omega).$$

Из ф-л (6.50) и (6.52) имеем

$$S_{nto}(\omega) = 2\,k\,T\,|H\,(\omega)|^2\left[R_{\text{экв}} + R_g\left(\frac{R_s}{R_s+R_g}\right)^2 + R_s\left(\frac{R_g}{R_s+R_g}\right)^2\right] =$$

$$= 2\cdot 1{,}38\cdot 10^{-23}\cdot 290\cdot\frac{1{,}56\cdot 10^{12}}{10^{10}+\omega^2}\cdot 2010 = \frac{2{,}5\cdot 10^{-5}}{10^{10}+\omega^2}.$$

Теперь средняя мощность шумового напряжения на выходе усилителя

$$\overline{u_o^2} = \frac{1}{\pi}\int_0^\infty\frac{2{,}5\cdot 10^{-5}}{10^{10}+\omega^2}d\,\omega = \frac{2{,}5\cdot 10^{-5}}{\pi}\cdot\frac{1}{10^5}\,\text{arc tg}\,\frac{\omega}{10^5}\Big|_0^\infty = 1{,}25\cdot 10^{-10}.$$

Следовательно, эффективное значение выходного шума равно $\sqrt{\overline{u_0^2}} = 11{,}2$ *мкв.*

Пример 6.5.

Определим оптимальное сопротивление источника R_s для триодного усилителя из примера 6.4 и вычислим соответствующий ему коэффициент шума. Последний определяется соотношением (6.51). Вообще $R_s \ll R_g$ и ф-ла (6.51) принимает вид

$$F \approx 1 + \frac{R_{\text{экв}}}{R_s} + \frac{R_s}{R_g}. \tag{6.53}$$

Оптимальную величину R_s можно получить, если приравнять $\frac{dF}{dR_s} = 0$, т. е.

$$\frac{dF}{dR_s} = \frac{-R_{\text{экв}}}{R_s^2} + \frac{1}{R_g} = 0$$

или

$$(R_s)_{\text{опт}} = \sqrt{R_{\text{экв}}R_g} = (10^3\cdot 10^5)^{1/2} = 10\ \textit{ком}.$$

Из ф-лы (6.53) получаем коэффициент шума для $(R_s)_{\text{опт}}$

$$F = 1 + 0{,}1 + 0{,}1 = 1{,}2.$$

Если сопротивление источника равно 1000 *ом,* то можно реализовать оптимальный коэффициент шума, применив повышающий трансформатор с коэффициентом трансформации $1 : \sqrt{10}$. Заметим, что для получения оптимального сопротивления источника нежелательно включать последовательно с ним дополнительное сопротивление 9000 *ом.* Это просто приведет к увеличению шума источника по сравнению с шумом усилителя.

Коэффициент шума, как уже указывалось, представляет собой отношение полной спектральной плотности выходной переменной к спектральной плотности мощности шума на выходе, обусловленной только входным источником. Может возникнуть вопрос относительно того, какую выходную переменную следует использовать:

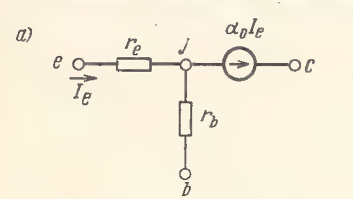
ток или напряжение. В действительности коэффициент шума не зависит от выходной переменной. Объясняется это тем, что напряжение и ток связаны между собой через сопротивление нагрузки, а так как коэффициент шума является функцией отношения спектров плотности мощности сигнала и шума, то зависимость от сопротивления нагрузки исключается.

Пример 6.6.

Определим коэффициент шума транзисторного усилителя, включенного по схеме с общей базой. Транзистор имеет в основном три вида шума: дробовой, шум токораспределения и тепловой. Дробовой шум можно представить с помощью источника тока $i_{sh}(t)$, подключенного к клеммам ej (рис. 6.19), со спектром плотности мощности (на низких частотах [1])):

$$S_{sh}(\omega) = q\, I_e, \tag{6.54а}$$

где I_e — среднее значение (постоянная составляющая) тока эмиттера.



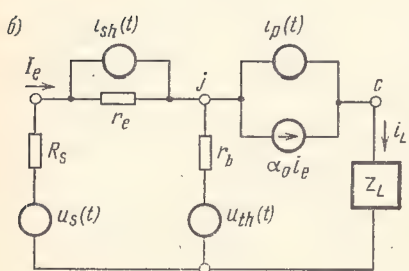


Рис. 6.19. К определению коэффициента шума транзисторного усилителя: *а*) эквивалентная схема транзистора; *б*) схема транзистора с источниками шума

Шум токораспределения можно представить источником тока $i_p(t)$, подключенным к клеммам jc, то спектром плотности мощности (на низких частотах)

$$S_p(\omega) = q\, \alpha_0 I_e (1 - \alpha_0), \tag{6.54б}$$

где α_0 — коэффициент усиления по постоянному току транзисторного усилителя, включенного по схеме с общей базой.

Причиной возникновения теплового шума является сопротивление $u_{th}(t)$ в цепи базы. Шум имеет спектр плотности мощности

$$S_{th}(\omega) = 2kT r_b, \tag{6.54в}$$

где r_b — распределенное сопротивление базы.

На рис. 6.19*а* приведена эквивалентная схема транзистора, а на рис. 6.19*б* показана эквивалентная схема с источниками шума и соответствующими сопротивлениями. Сопротивление входного источника представлено нешумящим резистором R_s, включенным последовательно с источником теплового шума $u_s(t)$, который имеет спектр плотности мощности

$$S_{R_s}(\omega) = 2kT R_s. \tag{6.54г}$$

Шумы транзистора имеют три составляющие: $i_{sh}(t)$, $i_p(t)$ и $u_{th}(t)$. Оставим читателю возможность показать, что передаточные функции, связывающие ток нагрузки i_L с токами источников соответственно:

[1]) A Van der Ziel. Theory of shot Noise in Junction diodes and Junction transistors. «Proc. IRE», v. 43, 11, 1639—1646 (Nov. 1955).

$$H_{ish}(\omega) = \frac{\alpha_0 r_e}{R_s + r_b(1-\alpha_0) + r_e}, \qquad (6.55а)$$

$$H_{ip}(\omega) = \frac{r_b + r_e + R_s}{R_s + r_b(1-\alpha_0) + r_e}, \qquad (6.55б)$$

$$H_{th}(\omega) = \frac{\alpha_0}{R_s + (1-\alpha_0) + r_e}. \qquad (6.55в)$$

Подобным образом передаточная функция, связывающая выходной ток i_L с напряжением внешнего источника $u_s(t)$,

$$H_{us}(\omega) = \frac{\alpha_0}{R_s + r_b(1-\alpha_0) + r_e}. \qquad (6.55г)$$

По определению

$$F = 1 + \frac{S_{nao}(\omega)}{S_{nso}(\omega)}, \qquad (6.56)$$

где

$$S_{nao}(\omega) = S_{sh}(\omega)\,|H_{ish}^2 + S_p(\omega)\,|H_{ip}|^2 + S_{th}(\omega)\,|H_{uth}|^2 \qquad (6.57а)$$

и

$$S_{nso}(\omega) = S_{R_s}(\omega)\,|H_{u_s}|^2. \qquad (6.57б)$$

Подставив ф-лы (6.57), (6.54) и (6.55) в ф-лу (6.56), получим

$$F = 1 + \frac{r_b}{R_s} + \frac{q I_e}{2\kappa T R_s}\left[r_e^2 + \frac{1-\alpha_0}{\alpha_0}(r_b + r_e + R_s)^2\right]. \qquad (6.58)$$

В транзисторе динамическое сопротивление эмиттера связано с током эмиттера соотношением

$$r_e = kT/(q I_e). \qquad (6.59)$$

Подставив ф-лу (6.59) в (6.58), получим

$$F = 1 + \frac{r_b + r_e/2}{R_s} + \frac{(r_b + r_e + R_s)^2(1-\alpha_0)}{2\alpha_0 r_e R_s}. \qquad (6.60)$$

Заметим, что существует оптимальное значение R_s, при котором коэффициент шума минимален.

Полученный результат справедлив для частот $\omega < \sqrt{1-\alpha_0 \omega_\alpha}$, где ω_α — предельная частота усиления по току для транзистора. На низких частотах (ниже 1 *кгц*) основную роль играет фликкерный шум, поэтому коэффициент шума, определяемый выражениями (6.58) и (6.60), справедлив для средних частот. На высоких частотах увеличивается шум токораспределения. Общее выражение для спектра плотности мощности шума токораспределения (шумовой источник тока) имеет вид

$$S_p(\omega) = q\,\alpha_0 I_e(1-\alpha_0)\,\frac{1 + \left[\omega/\omega_\alpha \sqrt{(1-\alpha_0)}\right]^2}{1 + (\omega/\omega_\alpha)^2}.$$

На низких частотах это выражение переходит в (6.54б). Подставив последнее выражение в ф-лу (6.57), получим

$$F = 1 + \frac{r_b + \frac{r_e}{2}}{R_e} + \frac{(r_b + r_e + R_s)^2(1-\alpha_0)\left[1 + \frac{1}{1-\alpha_0}\left(\frac{\omega}{\omega_\alpha}\right)^2\right]}{2\alpha_0 r_e R_s}.$$

На рис. 6.20 показано изменение коэффициента шума в зависимости от частоты.

Коэффициент шума усилителя, включенного по схеме с общим эмиттером, вычисляется аналогичным образом. Можно показать, что этот коэффициент определяется тем же выражением, что и для схемы с общей базой[1]).

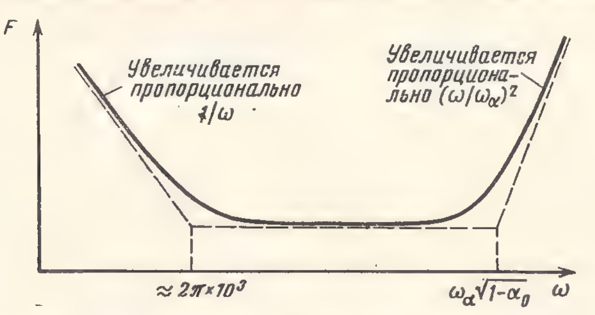


Рис. 6.20. Зависимость коэффициента шума транзисторного усилителя от частоты

6.7. Экспериментальное определение коэффициента шума

Экспериментально коэффициент шума усилителя можно определить с помощью установки, изображенной на рис. 6.21. Источником шума усилителя служит резистор R_s. Ко входу усилителя подключен диод, как это показано на рисунке. Диод работает в режиме насыщения, поэтому его динамическая проводимость равна нулю (бесконечное динамическое сопротивление). Сначала измеряется средняя мощность выходного напряжения $u_{no}(t)$ при отключенном диоде. Показание дает полную мощность шума на выходе N_{to}. Затем диод подсоединяется ко входу усилителя. Диод генерирует ток дробового шума со спектром плотности мощности qI_d, где I_d — постоянная составляющая тока диода. Средняя мощность выходного напряжения при этом увеличивается. Напряжение накала U_d подбирается таким, чтобы средняя мощность выходного шумового напряжения равнялась $2N_{to}$, и измеряется ток диода I_d. Можно легко показать, что средний коэффициент шума усилителя

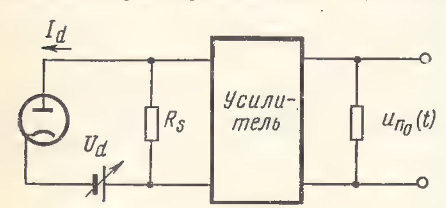


Рис. 6.21. К экспериментальному определению коэффициента шума усилителя

$$\overline{F} = qI_dR_s/(2kT). \tag{6.61}$$

[1]) E. G. Nielsen. Behavior of Noise Figure in Junction Transistors. «Proc. IRE», v. 45, № 7, pp. 957—963.

Справедливость ф-лы (6.61) можно доказать следующим образом. Диод имеет нулевую динамическую проводимость (поскольку работает в режиме насыщения, см. рис. 5.5) и поэтому представляет собой разомкнутую цепь. Следовательно, его можно заменить шумовым источником тока со спектром плотности мощности $S_d(\omega)$, равным [см. ф-лы (6.6) и (6.14)] $S_d(\omega) = qI_d$. Очевидно, включение шумового источника на входе усилителя увеличивает мощность выходного шума до значения N_{to}. Таким образом, мощность шума на выходе, обусловленная только источником входного шума с плотностью мощности qI_d, равна N_{to}.

В усилителе (без диода) источником шума на входе является резистор R_s, который можно представить нешумящим сопротивлением R_s, включенным параллельно источнику шумового тока (рис. 6.9*в*) с плотностью мощности $2kTG_s$ $(G_s = 1/R_s)$. Если параллельный источник тока с плотностью мощности qI_d увеличивает мощность шума в нагрузке до значения N_{to}, то аналогичный источник с плотностью мощности $2kTG_s$ приведет к увеличению мощности шума в нагрузке, равному $2kTG_s$ (N_{to}/qI_d). Но по определению это есть N_{so} — мощность шума в нагрузке, обусловленная только источником шума $N_{so} = 2kTG_sN_{to}/(qI_d)$.

Таким образом, средний коэффициент шума $\overline{F}$:

$$\overline{F} = N_{to}/N_{so} = qI_d/(2kTG_s) = qI_dR_s/(2kT).$$

Заметим, что этим способом определяется среднее значение коэффициента шума $\overline{F}$. Если же требуется получить спектральный коэффициент шума F, необходимо выполнять измерения во всем диапазоне частот, разделив его на узкие полосы (с помощью перестраиваемого узкополосного фильтра на выходе усилителя). Теперь измерение (6.61) дает величину F на центральной частоте фильтра. Для того чтобы определить значение F во всем диапазоне частот, узкополосный фильтр перестраивают по частоте, проходя весь частотный диапазон.

6.8. Плотность мощности шума

ОБЩИЕ ПОЛОЖЕНИЯ

До сих пор рассматривались нормализованные значения мощности сигнала и плотности мощности шума на сопротивлении в 1 *ом*. Мощность сигнала P определялась как мощность, рассеиваемая напряжением $f(t)$ на сопротивлении 1 *ом*. Если это же напряжение $f(t)$ приложено к сопротивлению R, *ом*, очевидно будет рассеиваться другая мощность. Для того чтобы различать эти мощности, обозначим нормализованную мощность (на сопротивлении 1 *ом*) через P_n, а мощность, рассеиваемую сопротивлением R, через P_R. Легко видеть, что

$$P_R = P_n/R. \tag{6.62}$$

Таким образом, действительная мощность, выделяемая сигналом $f(t)$ на сопротивлении R, равна $1/R$ части P_n, определяемой ф-лой (2.21а).

То, что справедливо для мощности, справедливо и для спектра плотности мощности. Плотность мощности представляет собой мощность, рассеиваемую в единичной полосе частотными составляющими $f(t)$ на сопротивлении 1 *ом*. Следовательно, плотность мощности является нормализованной. Если этот же сигнал $f(t)$ приложен к сопротивлению R, действительная плотность мощности равна $1/R$ части плотности мощности $f(t)$

$$S_R(\omega) = S_n(\omega)/R, \tag{6.63}$$

где $S_n(\omega)$ — плотность мощности сигнала $f(t)$, определяемая по ф-ле (2.22); $S_R(\omega)$ — действительная плотность мощности, рассеиваемая сигналом $f(t)$ на сопротивлении R *ом*.

НОМИНАЛЬНАЯ ПЛОТНОСТЬ МОЩНОСТИ ШУМА

Рассмотрим источник напряжения $f(t)$ с плотностью мощности $S_f(\omega)$ и полным внутренним сопротивлением R_s+iX_s. Как известно, для получения максимальной мощности на нагрузке необходимо, чтобы полное сопротивление нагрузки было комплексно-сопряженным с полным сопротивлением источника (согласованная нагрузка): $Z_L = Z_s^* = R_s - iX_s$.

На рис. 6.22 показано включение согласованной нагрузки. Очевидно, что в случае согласования напряжение $f(t)$ приложено к сопротивлению $2R_s$ и действительная плотность мощности, выделяемая $f(t)$, равна $S_f(\omega)/(2R_s)$. Половина этой мощности теряется на сопротивлении источника, а другая половина на нагрузке. Следовательно, на нагрузке рассеивается плотность мощности $S_f(\omega)/(4R_s)$. Эта максимальная плотность мощности, поступающей в нагрузку, называется номинальной и обозначается как

$$S_{av}(\omega) = S_f(\omega)/(4R_s). \tag{6.64}$$

Заметим еще раз, что номинальная плотность мощности является действительной плотностью мощности, которую можно получить от данного источника. Для источника напряжения $f(t)$ с полным внутренним сопротивлением R_s+iX_s номинальная плотность мощности равна $1/(4R_s)$ части нормализованной плотности мощности напряжения $f(t)$. В общем случае R_s зависит от частоты ω и ф-лу (6.64) следует записать в виде

$$S_{av}(\omega) = S_f(\omega)/[4R_s(\omega)]. \tag{6.65}$$

Нетрудно показать, что для источника тока $i(t)$ с плотностью мощности $S_i(\omega)$ и внутренней полной проводимостью $Y = G_s(\omega) + iB_s(\omega)$ номинальная плотность мощности $S_{av}(\omega)$ определяется выражением

$$S_{av}(\omega) = S_i(\omega)/(4G_s). \tag{6.66}$$

Найдем номинальную плотность мощности теплового шума пассивной цепи, содержащей только элементы R, L и С. Такую цепь (рис. 6.23) можно представить эквивалентным полным со-

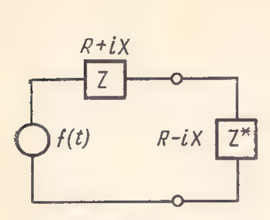


Рис. 6.22. Источник напряжения с согласованной нагрузкой

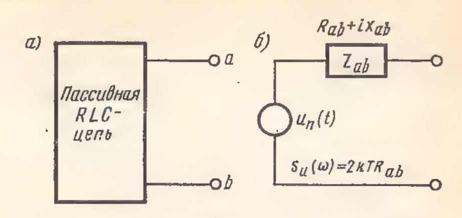


Рис. 6.23. К определению номинальной плотности мощности шума в пассивной цепи:
а) *R*, *L*, С-цепь; *б*) источник шумового напряжения с эквивалентным полным сопротивлением

противлением $Z_{ab}=R_{ab}(\omega)+iX_{ab}(\omega)$. На основании обобщенной теоремы Найквиста [ф-ла (6.34б) и рис. 6.10б] тепловой шум можно представить источником напряжения с плотностью мощности $2kTR_{ab}(\omega)$ (рис. 6.23б). Для получения наибольшей мощности шума в этой цепи необходимо к клеммам ab подключить нагрузку R_{ab}—iX_{ab}. Номинальная плотность мощности такого источника определится тогда в соответствии с ф-лой (6.65) как

$$S_{av}(\omega)=2kTR_{ab}(\omega)/[4R_{ab}(\omega)]=kT/2. \qquad (6.67)$$

Таким образом, номинальная плотность мощности шума (теплового) в любой пассивной R, L, C-цепи постоянна и равна $kT/2$ (предполагается, что все резисторы цепи имеют одинаковую температуру T).

Заметим, что максимальная мощность шума, которую можно получить от R, L, C-цепи в полосе частот Δf, находится как $2\Delta f(kT/2)=kT\Delta f$.

Следовательно, величина kT представляет собой номинальную мощность шума (теплового) в единичной полосе для любого R, L, C-двухполюсника.

6.9. Эффективная шумовая температура

Выше было получено, что номинальная плотность мощности шума R, L, C-двухполюсника равна $kT/2$. Этот результат не относится к двухполюснику, содержащему иные, чем тепловой, источники шума, например источник дробового шума. Тем не менее его можно распространить на все двухполюсники, введя в рассмотрение эффективную шумовую температуру T_n. Если номинальная плотность мощности шума произвольного двухполюсника равна

$S_{av}(\omega)$, то эффективную шумовую температуру T_n можно определить как

$$S_{av}(\omega) = kT_n/2 \qquad (6.68)$$

или

$$T_n = 2S_{av}(\omega)/k. \qquad (6.69)$$

Согласно ф-ле (6.69) в случае постоянства $S_{av}(\omega)$ в интересующем нас диапазоне частот эффективная шумовая температура T_n также постоянна. Если же, однако, $S_{av}(\omega)$ изменяется с частотой, T_n также изменяется с частотой ω. Номинальная мощность шума в полосе Δf любого двухполюсника равна $kT_n\Delta f$. Заметим, что для цепи, содержащей только R, L, C-элементы, эффективная шумовая температура T_n равна окружающей температуре T.

6.10. Выражение коэффициента шума через номинальный коэффициент усиления

Как уже известно, коэффициент шума можно выразить отношением спектров плотности мощности сигнала и шума на входе, деленным на аналогичное отношение на выходе [ф-ла (6.47)]. Можно также выразить коэффициент шума через номинальные плотности мощности и номинальный коэффициент усиления по мощности. Рассмотрим, например, схему, изображенную на рис. 6.24. На этом рисунке $S_s(\omega)$ и $S_n(\omega)$ представляют собой спектры плотности мощности напряжений сигнала и шума соответственно. Через $S_n(\omega)$ обозначен спектр теплового шума источника с полным сопротивлением $Z_s = R_s + iX_s$, и, следовательно, сопротивление $Z_s(\omega)$ на рис. 6.24 является нешумящим. Отношение спектров плотности мощности на клеммах ab обозначим через

$$S/N = S_s(\omega)/S_n(\omega). \qquad (6.70)$$

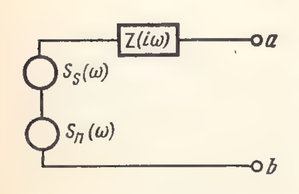


Рис. 6.24. Источник сигнала с шумом

Вычислим затем отношение номинальных спектров плотности мощности на клеммах ab. Так как $s(t)$ и $n(t)$ независимы, номинальную плотность мощности, создаваемую каждым источником, можно вычислить, предполагая, что другие источники отсутствуют. Таким образом, номинальная плотность мощности сигнала на клеммах ab равна $S_s(\omega)/(4R_s)$. Аналогично номинальная плотность мощности шума равна $S_n(\omega)/(4R_s)$. Следовательно, отношение номинальных плотностей мощности сигнала и шума на клеммах ab

$$(S/N)_{av} = S_s(\omega)/S_n(\omega). \qquad (6.71)$$

Полученное выражение показывает, что отношение плотностей мощности сигнала и шума в любых точках схемы равно отноше-

нию номинальных плотностей мощности сигнала и шума в тех же точках. Этот результат в действительности оказывается весьма общим. Можно легко показать, что отношение плотностей мощности сигнала и шума в любых точках равно действительному отношению плотностей мощности сигнала и шума, рассеиваемых на любой нагрузке в этих же точках. Выражение (6.71) является частным случаем общего результата для случая согласования полных сопротивлений нагрузки и источника.

Этот результат можно использовать для нахождения коэффициента шума усилителя. Согласно ф-ле (6.47) коэффициент шума усилителя равен отношению спектров плотности мощности сигнала и шума на входе, деленному на такое же отношение на выходе. Из ф-лы (6.71) следует, что коэффициент шума F можно также найти как отношение номинальных плотностей мощности сигнала и шума на входе, деленное на аналогичное отношение на выходе. Таким образом, $(S_{si})_{av}$ является номинальной плотностью мощности на входе. В соответствии с ф-лой (6.47) можно записать

$$F = \frac{(S_{si})_{av}/(S_{nsi})_{av}}{(S_{so})_{av}/(S_{nto})_{av}} = \frac{(S_{si})_{av}\,(S_{nto})_{av}}{(S_{so})_{av}\,(S_{nsi})_{av}}. \tag{6.72}$$

Определим номинальный коэффициент усиления мощности G усилителя как

$$G = (S_{so})_{av}/(S_{si})_{av}. \tag{6.73}$$

Заметим, что в общем случае номинальный коэффициент усиления мощности G зависит от частоты и, следовательно, может быть записан как $G(\omega)$. Подставив ф-лу (6.73) в (6.72), получим

$$F = (S_{nto})_{av}/[G\,(S_{nsi})_{av}], \tag{6.74}$$

где $(S_{nto})_{av}$ — номинальная плотность мощности шума на выходе, $(S_{nsi})_{av}$ — номинальная плотность мощности шумового источника на входе.

Если источник является пассивным R, L, C-двухполюсником, то согласно ф-ле (6.67) имеем

$$(S_{nsi})_{av} = kT/2. \tag{6.75}$$

Если же источник возбуждает шум не теплового происхождения, то

$$(S_{nsi})_{av} = kT_n/2. \tag{6.76}$$

В случае источника в виде пассивной R, L, C-цепи соотношение (6.74) принимает вид

$$F = 2\,(S_{nto})_{av}/(GkT). \tag{6.77}$$

Для источника, генерирующего шум, отличный от теплового, температуру T в ф-ле (6.77) следует заменить эффективной шумовой температурой T_n. Из ф-лы (6.77) получаем

$$(S_{nto})_{av} = FGkT/2. \tag{6.78}$$

Таким образом, номинальный спектр плотности мощности шума на выходе равен $FGkT/2$. Заметим, что $(S_{nto})_{av}$ складывается из двух составляющих:

$$(S_{nto})_{av} = (S_{nso})_{av} + (S_{nao})_{av}, \qquad (6.79)$$

где $(S_{nso})_{av}$ и $(S_{nao})_{av}$ — номинальные спектры плотности мощности шума на выходе, обусловленные шумом источника и собственными шумами усилителя соответственно.

Из ф-лы (6.73) следует, что

$$(S_{nso})_{av} = G(S_{nsi})_{av} = GkT/2. \qquad (6.80)$$

Поэтому на основании ф-л (6.78) — (6.80)

$$(S_{nao})_{av} = (F-1)\,GkT/2. \qquad (6.81)$$

В результате номинальный спектр плотности мощности на выходе, обусловленный собственными шумами усилителя, равен $(F-1)(GkT/2)$.

6.11. Коэффициент шума многокаскадного усилителя

Коэффициент шума многокаскадного усилителя можно выразить через коэффициенты шума отдельных каскадов. Наиболее существенным при определении общего коэффициента шума усилителя является коэффициент шума первого каскада, так как шум предыдущих каскадов усиливается последующим. Рассмотрим, например, двухкаскадный усилитель, изображенный на рис. 6.25. Для определения общего коэффициента шума усилителя F_{ab} необходимо сначала определить $(S_{nto})_{av}$ — полный номинальный спектр плотности мощности шума на его выходе, которая состоит из двух слагаемых: S_1 — номинального спектра плотности мощности шума на выходе, обусловленного первым каскадом; S_2 — номинального спектра плотности мощности шума на выходе, обусловленного шумом второго каскада. Составляющая S_1, очевидно, в G_b раз больше номинального спектра плотности мощности шума на выходе второго каскада. Ввиду того, что номинальная мощность шума на входе второго каскада равна полной номинальной мощности на выходе первого каскада и согласно ф-ле (6.78), имеем

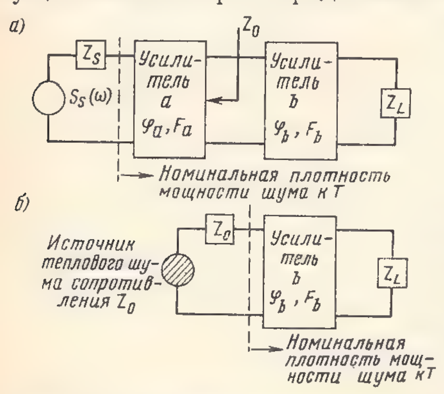


Рис. 6.25. К определению коэффициента шума двухкаскадного усилителя: *а)* последовательное включение двух усилителей; *б)* эквивалентная схема с заменой первого усилителя источником шума

$$S_1 = G_b(F_a G_a kT/2). \qquad (6.82)$$

Обозначим через Z_0 полное выходное сопротивление первого каскада, являющееся сопротивлением источника для второго каскада. Так как величина S_2 обусловлена собственным шумом только второго каскада усилителя, физическая природа эквивалентного источника несущественна. Предположим, что Z_0 является полным сопротивлением источника теплового шума и поэтому его номинальная плотность мощности равна $kT/2$. Заметим, что это предположение несущественно для вычисления S_2.

Пусть коэффициент шума усилителя b (рис. 6.25) равен F_b. Номинальная плотность мощности шума на выходе, обусловленная только усилителем b,

$$S_2 = (F_b - 1) G_b kT/2. \tag{6.83}$$

Следовательно,

$$(S_{nto})_{av} = S_1 + S_2 = \frac{kT}{2} [(F_b - 1) G_b + F_a G_a G_b]. \tag{6.84}$$

Но из ф-лы (6.78) имеем

$$(S_{nto})_{av} = F_{ab} G_{ab} kT/2, \tag{6.85}$$

где G_{ab} — номинальный коэффициент усиления двухкаскадного усилителя.

Из определения номинального коэффициента усиления легко заметить, что усиление двухкаскадного усилителя равно произведению коэффициентов отдельных каскадов $G_{ab} = G_a G_b$. Следовательно,

$$(S_{nto})_{av} = F_{ab} G_a G_b kT/2. \tag{6.86}$$

Сравнивая ф-лы (6.84) и (6.86), получаем

$$F_{ab} = F_a + \frac{F_b - 1}{G_a}. \tag{6.87}$$

Напомним, что F_b — коэффициент шума усилителя b при условии, что его входным шумовым источником является пассивное полное сопротивление Z_0, равное выходному полному сопротивлению усилителя a.

В общем случае многокаскадного усилителя

$$F = F_a + \frac{F_b - 1}{G_a} + \frac{F_c - 1}{G_a G_b} + \ldots \tag{6.88}$$

Из ф-лы (6.88) следует, что наиболее существенным с точки зрения коэффициента шума является первый каскад усилителя. Поэтому для того чтобы коэффициент шума усилителя был минимальным, наименее шумящим должен быть первый каскад.

Пример 6.7.

Определим коэффициент шума двухкаскадного усилителя, схема которого приведена на рис. 6.26. В усилителе применены одинаковые лампы с пара-

метрами $g_m=2{,}5\cdot10^{-3}$ *сим* и $r_i=10\,000$ *ом*. В рабочем диапазоне частот блокирующую емкость можно заменить коротким замыканием. Для удобства разделим усилитель на два каскада, как это показано пунктиром на рис. 6.26, и

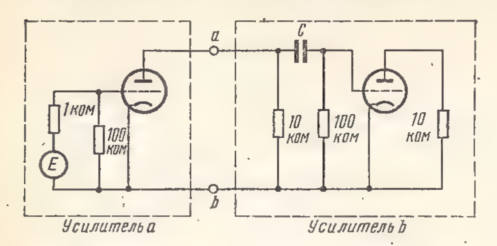


Рис. 6.26. Принципиальная схема двухкаскадного усилителя

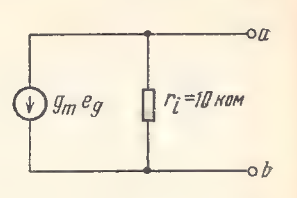


Рис. 6.27. Эквивалентная выходная цепь усилителя

найдем коэффициенты шума F_a и F_b обоих каскадов, пользуясь ф-лой (6.51).

Для усилителя *a* $R_{экв}=2{,}5/(2{,}5\cdot10^{-3})=1000$ *ом*, $R_s=1000$ *ом* и $R_g=$ $=100\,000$ *ом*.

Подставив эти величины в ф-лу (6.51), получим $F_a=2{,}03$.

Для усилителя *b* $R_{экв}=2{,}5(2{,}5\cdot10^{-3})=1000$ *ом*, а $R_g=9100$ *ом* представляет собой параллельное соединение сопротивлений в 10 000 и 100 000 *ом*. Полным сопротивлением источника для усилителя *b* служит выходное сопротивление усилителя *a*, представляющее собой сопротивление анод-катод r_i первой лампы: $R_s=r_i=10\,000$ *ом*. Подставив эти величины в ф-лу (6.51), получим $F_b=2{,}54$.

Теперь вычислим номинальный коэффициент усиления мощности усилителя *a*, который определяется ф-лой (6.73). Заметим, что на любой частоте коэффициент усиления *G* можно также представить как

$$G=\frac{\text{номинальная мощность сигнала на нагрузке}}{\text{номинальная мощность сигнала на входе}}. \qquad (6.89)$$

Рассмотрим усилитель *a* с сопротивлением источника 1000 *ом*. Если напряжение сигнала равно *E в*, то номинальная мощность получается включением сопротивления нагрузки 1000 *ом*. Максимальная мощность, отдаваемая в нагрузку, находится как

$$(E/2000)^2\,1000=E^2/4000,\ вт. \qquad (6.90)$$

На рис. 6.27 показана эквивалентная цепь со стороны выходных клемм усилителя *a*. Величина тока эквивалентного источника равна $g_m e_g$, где e_g — напряжение на сетке первой лампы. Из схемы усилителя, приведенной на рис. 6.26, имеем

$$e_g=\frac{100}{101}\,E,\quad g_m e_g=2{,}5\cdot10^{-3}\cdot\frac{100}{101}\,E\approx2{,}5\cdot10^{-3}E.$$

Максимальная мощность на клеммах нагрузки *ab* равна мощности, выделяемой на согласованной нагрузке. Это получится, если сопротивление нагрузки на клеммах *ab* будет равно 10 000 *ом* и мощность на выходе при этом

$$(g_m e_g/2)\,10\,000=1{,}56\cdot10^{-2}\,E^2. \qquad (6.91)$$

Из ф-л (6.90) и (6.91) получаем

$$G_a=\frac{1{,}56\cdot10^{-2}}{E^2/4000}=62{,}4. \qquad (6.92)$$

Общий коэффициент шума

$$F_{ab} = F_a + \frac{F_b - 1}{G_a} = 2,03 + \frac{1,54}{62,4} = 2,03 + 0,024 = 2,0547 \approx F_a.$$

6.12. Каскодный усилитель

Из изложенного в § 6.11 следует, что в усилителе, состоящем из нескольких каскадов, общий коэффициент шума определяется в основном усилением первого каскада, причем общий коэффициент шума уменьшается при увеличении усиления первого каскада. В полосовом усилителе увеличение усиления может привести к самовозбуждению, поэтому усиление нельзя увеличивать выше определенного предела, не рискуя нарушить устойчивость усилителя. Внимательно рассматривая ф-лу (6.88), можно, однако, заметить, что G_a является номинальным коэффициентом усиления мощности, а не напряжения. Можно создать усилитель с большим номинальным усилением мощности и в то же время малым усилением напряжения. Такой первый каскад усилителя уменьшит общий коэффициент шума без опасности возбуждения. В этом состоит принцип каскодного усилителя, в котором первый каскад имеет очень большой коэффициент усиления мощности и вместе с тем очень малое усиление напряжения [1]).

ЗАДАЧИ

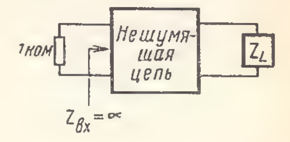
6.1. Выведите формулу эффективного значения шумового тока в RL-цепи.

6.2. Определите спектр плотности мощности шумового тока, протекающего через последовательную L, C, R-цепь.

6.3. Два резистора сопротивлением по 1000 *ом* каждый имеют температуру 300°K и 400°K соответственно. Определите спектр плотности мощности шумового напряжения в цепи, образованной последовательным соединением этих резисторов.

6.4. Повторите задачу 6.3 для параллельно включенных резисторов.

6.5. Определите эффективное значение напряжения шума на выходе нешумящей цепи, показанной на рисунке, если ко входу цепи подключен резистор сопротивлением 1 *ком* и ее передаточная функция по напряжению соответствует:



а) идеальному фильтру нижних частот с частотой среза f_c, *гц;*

б) идеальному полосовому фильтру с полосой пропускания f_c, *гц,* и центральной частотой f_0;

в) экспоненциальному фильтру $H(\omega) = Ae^{-|\omega|/\omega_0}$,

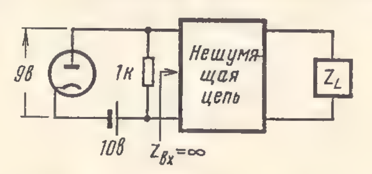
г) гауссову фильтру $H(\omega) = Ae^{-\omega^2/\omega_0^2}$

---

[1]) См. T. M. Pettit and M. M. Mc Whorter. Electonic Amplifier Circuits. «Mc Graw-Hill, New Jork, 1961.

Предполагается, что входное сопротивление цепи равно бесконечности. Шумом, обусловленным нагрузкой Z_L, можно пренебречь.

6.6. Повторите задачу 6.5 при условии, что на входе цепи параллельно резистору сопротивлением 1 *ком* включен диод (см. рисунок). Предполагается,



что диод работает в режиме пространственного заряда с динамическим сопротивлением 6670 *ом* и температурой катода 1000°.

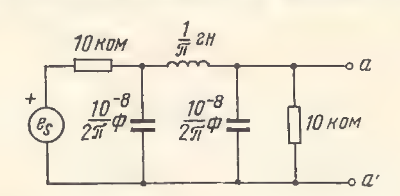
6.7. Определите спектр плотности мощности напряжения шума на клеммах aa' резистивной цепи, изображенной на рисунке, следующими способами:

а) вычисляя спектр плотности мощности на клеммах aa' как сумму спектров плотности мощности, обусловленных каждым из трех сопротивлений;

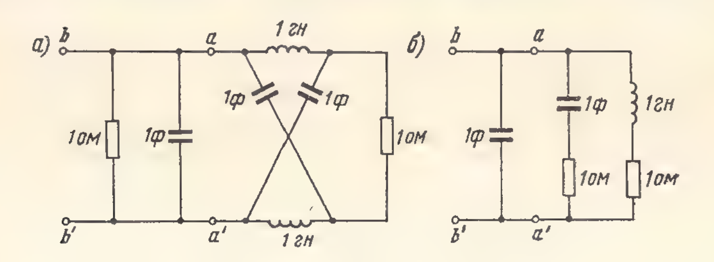
б) путем вычисления эквивалентного сопротивления $R_{aa'}$, в точках aa', а затем определения требуемой плотности мощности $2kTR_{aa'}$.

Определите спектр плотности мощности напряжения на клеммах aa', если резисторы R_1, R_2 и R_3 имеют различные температуры T_1, T_2 и T_3 соответственно.

6.8. Определите спектр плотности мощности и эффективное значение шумового напряжения на выходных клеммах aa' фильтра нижних частот, изображенного на рисунке. Этот фильтр известен как фильтр Баттерворса третьего порядка.



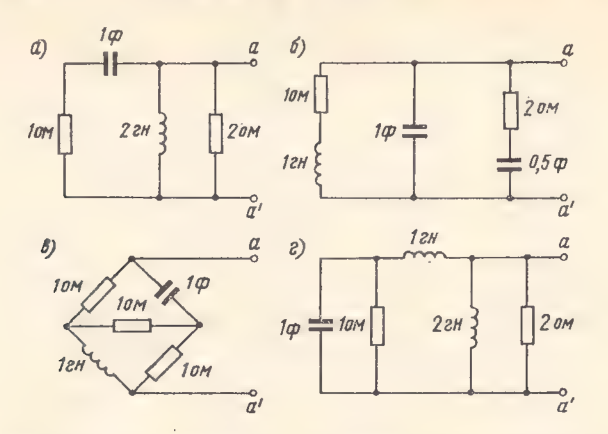
6.9. Найдите эффективное значение шумового напряжения на клеммах bb' для каждой из приведенных на рисунке цепей следующими способами:



а) путем вычисления спектра плотности мощности на клеммах bb' как суммы спектров плотности мощности, обусловленных каждым из резисторов;

б) путем замены цепи, расположенной правее точек aa', эквивалентной полной проводимостью и использования ф-лы (6.34).

6.10. Определите спектр плотности мощности шумового напряжения на клеммах aa' цепей, приведенных на рисунке.



6.11. Покажите, что эквивалентная шумовая полоса фильтра Баттерворса в задаче 6.8 относительно $f_0=0$ равна $10^4\ \pi/3$ *гц*.

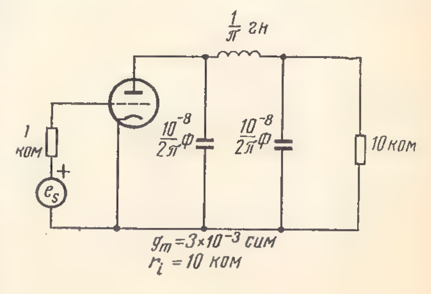
6.12. Для указанных ниже цепей вычислите эквивалентную шумовую полосу относительно нулевой частоты:

$$H_1(\omega) = A\,e^{-\alpha|\omega|},\quad H_2(\omega) = A\,e^{-\alpha\omega^2},\quad H_3(\omega) = S_a(\omega t_0),\quad H_4(\omega) = G_{2W}(\omega).$$

Вычислите эквивалентную шумовую полосу относительно частоты ω_0, если

$$H(\omega) = S_a[(\omega-\omega_0)t_0] + S_a[(\omega+\omega_0)t_0].$$

6.13. Фильтр Баттерворса третьего порядка, рассмотренный в задаче 6.8, подключен к выходным клеммам триодного усилителя (см. рисунок). Сопротивление источника 1000 *ом*. Определите эффективное значение шумового напряжения на нагрузке. Вычислите спектральный и средний коэффициенты шума усилителя. Известно, что $g_m = 3\cdot 10^{-3}$ *сим*, и $r_i = 10\,000$ *ом*.



6.14. Найдите прямым методом коэффициент шума двухкаскадного усилителя, показанного на рис. 6.25, используя ф-лу (6.44).

6.15. Покажите, что коэффициенты шума усилителей, выполненных на вакуумной лампе по схеме с заземленным катодом, заземленной сеткой и заземленным анодом, примерно одинаковы.

6.16. Покажите, что если эффективная шумовая температура источника постоянна и не зависит от частоты, то средний коэффициент шума $\overline{F}$ усилителя определяется формулой

$$\overline{F} = \int_0^\infty F\,G_a\,df \Big/ \int_0^\infty G_a\,df,$$

где G_a — номинальный коэффициент усиления мощности.

6.17. Определите коэффициент шума усилителя, рассмотренного в примере 6.4 (стр. 221), используя ф-лы (6.77) или (6.81).

ГЛАВА

7

Помехоустойчивость систем связи

7.1. Представление шума с полосовым спектром через квадратурные составляющие

В гл. 3 и 5 при рассмотрении различных способов модуляции отмечалось, что одни системы передачи более помехоустойчивы, другие — менее помехоустойчивы. Причем чем шире полоса частот сигнала, тем выше помехоустойчивость. До сих пор эти выводы основывались целиком на качественном анализе. Здесь же будут получены количественные соотношения между отношением сигнал/шум на выходе приемника (являющемся мерой помехоустойчивости системы передачи) и шириной спектра сигнала для различных способов передачи.

Однако прежде чем перейти к количественному анализу, рассмотрим один из способов представления шумов.

Как уже указывалось, после модуляции сигналы имеют полосовые спектры. Эти сигналы в процессе передачи искажаются широкополосным (обычно белым) шумом. Первая обработка сигнала в приемнике, очевидно, заключается в выделении той полосы частот, которую занимает сигнал, и отфильтровке всех частотных составляющих вне этой полосы. Выходное колебание такого полосового фильтра есть полезный сигнал (модулированный) плюс полосовой шум. Напряжение на выходе детектора состоит из полезного сигнала и шума, возникающего благодаря полосовому шуму на входе детектора. Поскольку в общем случае при вычислении выходного шума детектора приходится иметь дело с полосовым шумом, рассмотрим удобную форму представления последнего.

Пусть полосовой шум $n(t)$ имеет спектр плотности мощности, показанный на рис. 7.1. Покажем, что такое случайное шумовое колебание $n(t)$ можно записать в виде

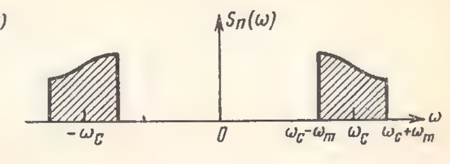
$$n(t) = n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t, \qquad (7.1)$$

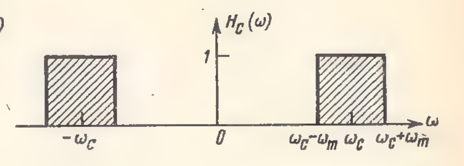
где $n_c(t)$ и $n_s(t)$ — низкочастотные случайные колебания, спектры которых ограничены частотой ω_m, *рад/сек*, причем мощности

(средние квадратические значения) колебаний $n(t)$, $n_c(t)$ и $n_s(t)$ одинаковы, т. е.

$$\overline{n^2(t)}=\overline{n_c^2(t)}=\overline{n_s^2(t)}. \qquad (7.2)$$

Для доказательства этого положения заметим, что колебание $n(t)$ при прохождении через идеальный полосовой фильтр, передаточная функция которого показана на рис. 7.1*б*, не искажается. Таким образом, колебание $n(t)$ можно передавать через такой фильтр без каких-либо изменений (рис. 7.1*в*).





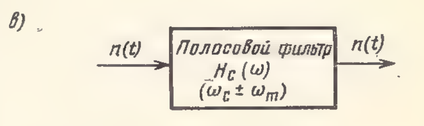


Рис. 7.1. К пояснению неискаженной передачи узкополосного случайного колебания через идеальный полосовой фильтр:

а) спектр входного процесса; *б*) передаточная функция фильтра; *в*) схема передачи

Полосовой фильтр с характеристикой $H_c(\omega)$ можно реализовать с помощью схемы, показанной на рис. 7.2*а*. Фильтр с характеристикой $H_0(\omega)$ (рис. 7.2*а*)—идеальный фильтр нижних частот с частотой среза ω_m, как показано на рис. 7.2*б*. Для доказательства этого приложим ко входу системы задержанный единичный импульс $\delta(t-\tau)$ (рис. 7.2*в*). Покажем, что на выходе получится сигнал $h_0(t-\tau)\cos\omega_c(t-\tau)$, где $h_0(t)$ — реакция фильтра нижних частот на единичный импульс.

Заметим, что

$$\delta(t-\tau)\cos\omega_c t=\delta(t-\tau)\cos\omega_c\tau,$$

$$\delta(t-\tau)\sin\omega_c t=\delta(t-\tau)\sin\omega_c\tau.$$

Поэтому для верхнего и нижнего фильтров нижних частот входными являются сигналы $2\delta(t-\tau)\cos\omega_c\tau$ и $2\delta(t-\tau)\sin\omega_c\tau$ соответственно. Если $h_0(t)$ обозначает импульсную реакцию любого из этих фильтров, то $h_0(t-\tau)$ есть отклик фильтров на сигнал $\delta(t-\tau)$. Очевидно, выходные сигналы фильтров равны $2h_0(t-\tau)\cos\omega_c\tau$ и $2h_0(t-\tau)\sin\omega_c\tau$. Далее они умножаются на $\cos\omega_c t$ и $\sin\omega_c t$ соответственно, а затем суммируются. Оконечный выходной сигнал

$$\varphi(\tau)=2h_0(t-\tau)(\cos\omega_c\tau\cos\omega_c t+\sin\omega_c\tau\sin\omega_c t)=$$

$$=2h_0(t-\tau)\cos\omega_c(t-\tau). \qquad (7.3)$$

Покажем теперь, что отклик идеального полосового фильтра

(рис. 7.1*б*) на единичный импульс $\delta(t-\tau)$ равен сигналу $\varphi(t)$, определяемому ф-лой (7.3). Из теоремы модуляции [ф-ла (1.116а)] следует, что если импульсная реакция $h_0(t)$ имеет спектр $H_0(\omega)$ (рис. 7.2*б*), то спектр $H_c(\omega)$ (рис. 7.1*б*) соответствует импульсной

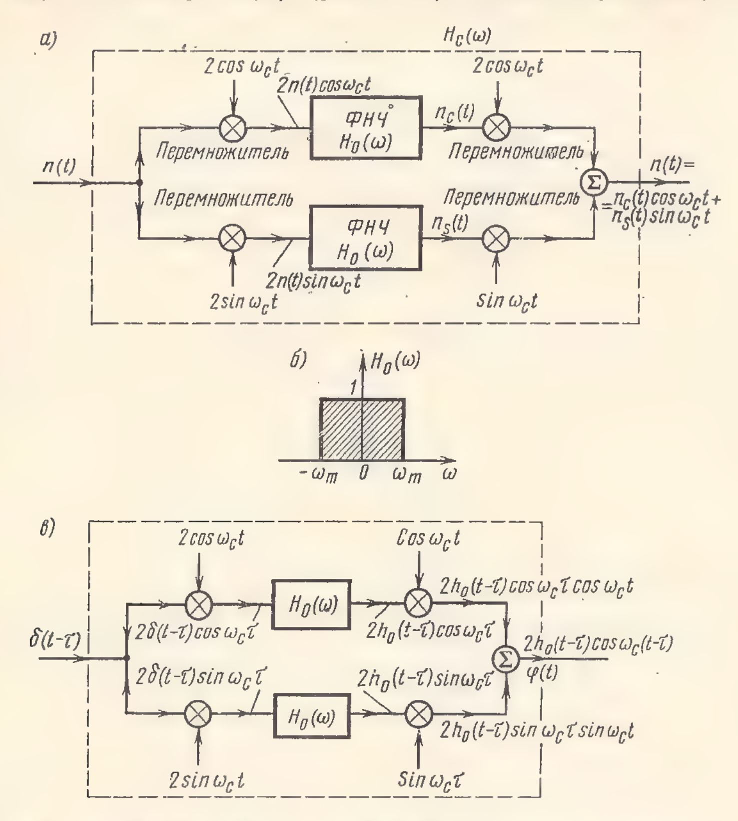


Рис. 7.2. Эквивалентные представления идеального полосового фильтра:
а) схема с выделением огибающих квадратурных составляющих и с последующим преобразованием частоты; *б*) передаточная функция *ФНЧ*; *в*) схема фильтра с указанием напряжений в различных точках

реакции $h_c(t)=2h_0(t)\cos\omega_c t$. Очевидно, реакция фильтра на $\delta(t-\tau)$ равна[1] $2h_0(t-\tau)\cos\omega_c(t-\tau)$. Это выражение совпадает с реакцией $\varphi(t)$ в ф-ле (7.3.) Следовательно, цепь рис. 7.2*а* действительно эквивалентна идеальному полосовому фильтру.

[1]) Здесь сравниваются реакции двух цепей на задержанный единичный импульс $\delta(t-\tau)$, а не на $\delta(t)$, как принято в случае систем с постоянными параметрами, поскольку цепь рис. 7.2*а* содержит элементы с переменными параметрами. Цепи являются эквивалентными, если их импульсные реакции одинаковы при любых значениях τ.

Если на вход идеального полосового фильтра подводится полосовой шум $n(t)$ (рис. 7.1*а*), то на выходе должен быть тоже шум $n(t)$. Обозначив выходные напряжения верхнего и нижнего полосовых фильтров на рис. 7.2*а* через $n_c(t)$ и $n_s(t)$ соответственно, можно записать

$$n(t) = n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t,$$

что совпадает с ф-лой (7.1).

Колебания $n_c(t)\cos\omega_c t$ и $n_s(t)\sin\omega_c t$ называются квадратурными составляющими процесса $n(t)$, а $n_c(t)$ и $n_s(t)$ — их огибающими. Определим далее спектры плотности мощности колебаний $n_c(t)$ и $n_s(t)$. Из рис. 7.2*а* видно, что $n_c(t)$ есть реакция фильтра нижних частот с характеристикой $H_0(\omega)$ на входной сигнал $2n(t)\cos\omega_c t$. Если $S_n(\omega)$ есть спектр плотности мощности колебания $n(t)$, то спектр плотности мощности колебания $n(t)\cos\omega_c t$ равен[1]) [см. ф-лу (2.24)]

$$\frac{1}{4}\left[S_n(\omega+\omega_c) + S_n(\omega-\omega_c)\right], \tag{7.4}$$

а спектр плотности мощности колебания $2n(t)\cos\omega_c t$

$$4\cdot\frac{1}{4}\left[S_n(\omega+\omega_c) + S_n(\omega-\omega_c)\right] = S_n(\omega+\omega_c) + S_n(\omega-\omega_c). \tag{7.5}$$

Спектры $S_n(\omega)$, $S_n(\omega+\omega_c)$, $S_n(\omega-\omega_c)$ и $[S_n(\omega+\omega_c)+S_n(\omega-\omega_c]$ показаны на рис. 7.3. Колебание $2n(t)\cos\omega_c t$, имеющее спектр $[S_n(\omega+\omega_c)+S_n(\omega-\omega_c)]$ (рис. 7.3*г*), пропускается через идеальный фильтр нижних частот $H_0(\omega)$ (рис. 7.2*а*), который подавляет все частотные составляющие с $|\omega|>\omega_m$. Получающийся спектр и есть спектр плотности мощности колебания $n_c(t)$, показанный на рис. 7.3*д*. Очевидно, $S_{n_c}(\omega)$ определяется выражением

$$S_{n_c}(\omega) = \begin{cases} S_n(\omega+\omega_c) + S_n(\omega-\omega_c) & \text{при } |\omega| < \omega_m, \\ 0 & \text{при } |\omega| > \omega_m. \end{cases} \tag{7.6}$$

Подобным же образом, используя ф-лу (2.24), можно показать, что спектр колебания $n_s(t)$ совпадает с $S_{n_c}(\omega)$:

$$S_{n_s}(\omega) = S_{n_c}(\omega). \tag{7.7}$$

---

[1]) Этот результат справедлив только в том случае, если $n(t)$—случайное колебание. Если $n(t)$ — не случайный сигнал, может появиться дополнительный спектр около частоты $\omega=0$.

Заметим, что если $n(t)$ — белый шум со спектральной плотностью $S_n(\omega)=N/2$, то

$$S_{n_c}(\omega)=S_{n_s}(\omega)=\begin{cases} N & \text{при } |\omega|<\omega_m, \\ 0 & \text{при } |\omega|>\omega_m. \end{cases} \qquad (7.8)$$

Далее определим средние мощности (средние квадратические значения) колебаний $n(t)$, $n_c(t)$ и $n_s(t)$. Мощность сигнала равна

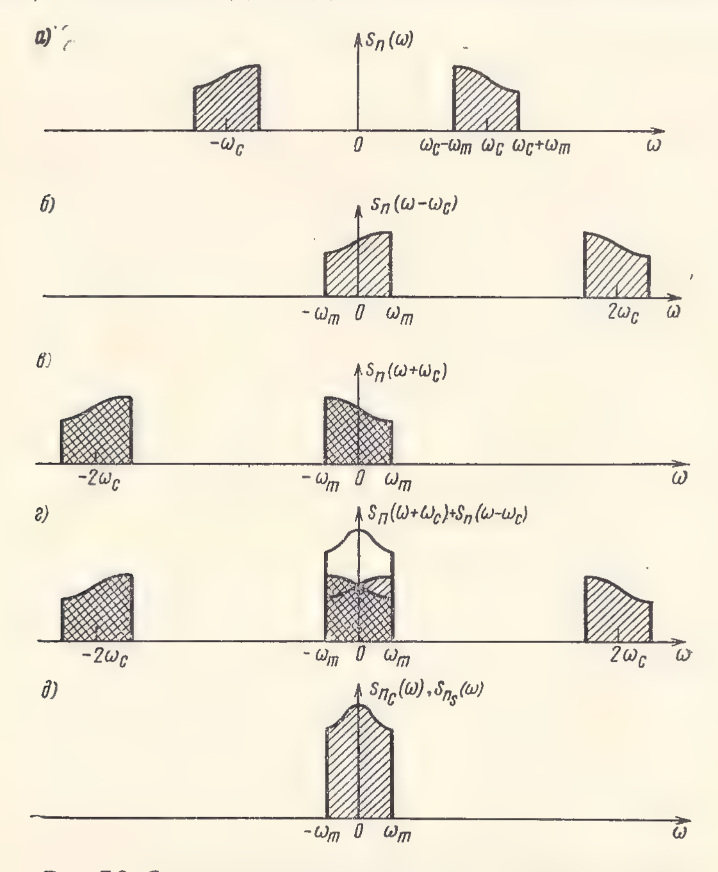


Рис. 7.3. Спектры плотности мощности в различных точках схемы рис. 7.2а:
а) спектр входного процесса; *б)* составляющая спектра $S_n(\omega-\omega_c)$; *в)* составляющая спектра $S_n(\omega+\omega_c)$; *г)* спектр на входе фильтра нижних частот; *д)* спектр на выходе фильтра нижних частот

площади под кривой его спектра плотности мощности, деленной на 2π. Из рис. 7.3 видно, что площадь под кривой спектра $S_n(\omega)$ равна площади под кривой спектра $S_{n_c}(\omega)$ или $S_{n_s}(\omega)$. Это объясняется тем, что спектр $S_{n_c}(\omega)$ образован суммированием спект-

ров $S_n(\omega)$ на положительных и отрицательных частотах. Отсюда и следует справедливость ф-лы (7.2) [1].

Заметим, что мощность шумовой составляющей $n_c(t)\cos\omega_c t$ равна $\overline{n_c^2}/2$, а составляющей $n_s(t)\sin\omega_c t$ — $\overline{n_s^2}/2$ [см. ф-лу (2.25)]. Следовательно, мощность шума делится поровну между двумя квадратурными составляющими.

Формулу (7.1) можно выразить в ином виде:

$$n(t) = R(t)\cos[\omega_c t + \theta(t)], \tag{7.9}$$

где

$$R(t) = \sqrt{n_c^2(t) + n_s^2(t)}, \tag{7.10}$$

$$\theta(t) = \operatorname{arc\,tg}\frac{n_s(t)}{n_c(t)}. \tag{7.11}$$

Поскольку $n_c(t)$ и $n_s(t)$ — медленно меняющиеся колебания, то из ф-л (7.10) и (7.11) следует, что $R(t)$ и $\theta(t)$ — также медленные процессы. Случайное колебание с узкополосным спектром имеет вид синусоидального колебания, амплитуда и фаза которого медленно изменяются во времени, как показано на рис. 7.4. Из ф-лы (7.9) вытекает, что огибающая этого колебания равна $R(t)$, а фаза — $\theta(t)$. Заметим, что такое колебание имеет как амплитудную, так и угловую модуляцию.

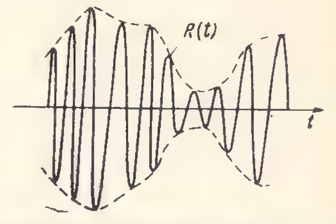


Рис. 7.4. Осциллограмма узкополосного случайного процесса

7.2. Общие замечания о помехоустойчивости систем связи

Перейдем к исследованию влияния вида модуляции на отношение сигнал/шум на выходе приемника. Ранее отмечалось, что широкополосные системы связи обычно менее чувствительны к воздействию шумов. Если сигнал преобразуется так, что его

---

[1]) Следует еще раз подчеркнуть, что эти результаты применимы только к полосовым случайным сигналам. Их нельзя без разбора применять к любым полосовым сигналам. Например, рассмотрим низкочастотный сигнал $f(t)$, спектр которого ограничен частотой ω_m. Тогда сигнал $\varphi(t) = f(t)\cos\omega_c t$ есть колебание с полосовым спектром. Если этот сигнал представляется ф-лой (7.1)

$$\varphi(t) = \varphi_c(t)\cos\omega_c t + \varphi_s(t)\sin\omega_c t,$$

то, очевидно, $\varphi_c(t) = f(t)$ и $\varphi_s(t) = 0$. Таким образом, ф-ла (7.1) несправедлива. Здесь сигнал $\varphi(t) = f(t)\cos\omega_c t$ не является стационарным случайным сигналом (даже если $f(t)$ — стационарный случайный сигнал). В таких случаях ф-ла (7.4) несправедлива. В общем случае, любой сигнал с полосовым спектром можно представить по ф-ле (7.1), но соотношение (7.2) справедливо только для стационарных случайных сигналов с полосовым спектром.

спектр расширяется, то после преобразования он становится более невосприимчивым к шумам. Иными словами, отношение сигнал/шум увеличивается. В гл. 8 на теоретико-информационной основе доказывается возможность обмена полосы частот на отношение сигнал/шум. Показывается, что данное количество информации можно передать, используя различные комбинации полосы частот и отношения сигнал/шум в канале. Если полоса частот уменьшается, то необходимо передавать сигнал бо́льшей мощности (т. е. требуется бо́льшее отношение сигнал/шум). С другой стороны, если для передачи доступна более широкая полоса частот, то же самое количество информации можно передать сигналом меньшей мощности. Ширина спектра сигнала зависит от вида применяемой модуляции. В этой главе будут рассмотрены соотношения между шириной спектра передаваемого сигнала и отношением сигнал/шум при различных видах модуляции.

Если бы шумы в канале передачи отсутствовали, то модулированный сигнал имел бы бесконечно большое отношение сигнал/шум на входе детектора. Однако, так как шумы в канале имеются, отношение сигнал/шум конечно. Детектор преобразует модулированный сигнал в модулирующий сигнал плюс шум. Модулированный сигнал на входе детектора и модулирующий на его выходе имеют различные полосы частот и отношения сигнал/шум. Следовательно, именно в детекторе выполняется обмен полосы частот на отношение сигнал/шум.

Ниже будет исследоваться природа обмена при различных видах модуляции в случае аддитивного шума, т. е. шума, взаимодействующего с сигналом посредством сложения. Соответствующая модель передачи показана на рис. 7.5. Заметим, что полосовой фильтр на входе приемника отфильтровывает лишь те шумы, которые лежат вне спектра сигнала. Таким образом, шум на входе детектора имеет ту же ширину спектра, что и модулированный сигнал.

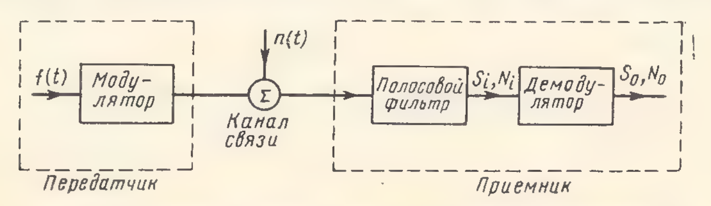


Рис. 7.5. Модель канала передачи с аддитивным шумом

Для каждой системы будет проводиться сравнение отношений сигнал/шум (по мощности) на входе и выходе детектора. Будем обозначать через S_i и S_o мощности полезного сигнала (его среднее квадратическое значение) на входе и выходе детектора соответственно, а через N_i и N_o — мощности шумов на входе и выходе детектора.

7.3. Помехоустойчивость систем связи с амплитудной модуляцией

СИСТЕМА ДБП-ПН

Схема приемника сигналов ДБП-ПН была приведена на рис. 3.1*е* и еще раз показана на рис. 7.6[1]). Пусть $f(t)$ есть сообщение, спектр которого ограничен частотой ω_m, *рад/сек* ($\omega_m \leqslant \omega_c$, где ω_c — несущая частота). Полезный сигнал на входе детектора

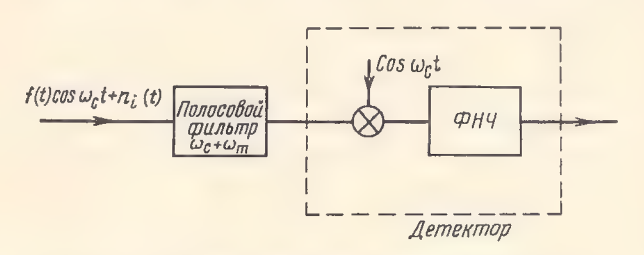


Рис. 7.6. Схема приемника сигналов ДБП-ПН

равен $f(t)\cos\omega_c t$. Ранее было показано [ф-ла (2.25б)], что средняя мощность сигнала $f(t)\cos\omega_c t$ равна половине средней мощности сообщения $f(t)$, т. е.

$$S_i = \overline{[f(t)\cos\omega_c t]^2} = \frac{1}{2}\overline{f^2(t)}. \tag{7.12}$$

Поскольку сигнал на выходе синхронного детектора в рассматриваемом случае равен $\frac{1}{2}f(t)$ [см. ф-лу (3.2)], то его средняя мощность

$$S_o = \overline{\left[\frac{1}{2}f(t)\right]^2} = \frac{1}{4}\overline{f^2(t)} = \frac{1}{2}S_i. \tag{7.13}$$

Чтобы вычислить мощности шумов на входе N_i и выходе N_o детектора, используем представление входного узкополосного шума в виде

$$n_i(t) = n_c(t)\cos\omega_c t + n_s(t)\sin\omega_c t, \tag{7.14}$$

где

$$N_i = \overline{n_i^2(t)}. \tag{7.15}$$

---

[1]) Полосовой фильтр на входе приемника в этом случае можно не включать, так как любые шумовые составляющие, лежащие вне полосы частот сигнала, подавляются фильтром нижних частот детектора.

Если шум $n_i(t)$ подводится ко входу синхронного детектора (который умножает входное колебание на $\cos\omega_c t$), то шум на выходе перемножителя

$$n_d(t) = n_i(t)\cos\omega_c t = n_c(t)\cos^2\omega_c t + n_s(t)\sin\omega_c t\cos\omega_c t =$$
$$= \frac{1}{2}\,[n_c(t) + n_c(t)\cos 2\omega_c t + n_s(t)\sin 2\omega_c t].$$

Слагаемые $n_c(t)\cos 2\omega_c t$ и $n_s(t)\sin 2\omega_c t$ имеют спектры колебаний $n_c(t)$ и $n_s(t)$, сдвинутые на $\pm 2\omega_c$; они отфильтровываются фильтром нижних частот на выходе детектора (см. рис. 7.6). Следовательно, шум на выходе детектора:

$$n_o(t) = \frac{1}{2}\,n_c(t), \qquad (7.16)$$

$$N_o = \overline{n_o^2(t)} = \frac{1}{4}\,\overline{n_c^2(t)}. \qquad (7.17)$$

Используя ф-лы (7.2) и (7.12), находим

$$N_o = \frac{1}{4}\,\overline{n_i^2(t)} = \frac{1}{4}\,N_i. \qquad (7.18)$$

Из ф-л (7.13) и (7.18) получаем

$$\frac{S_o/N_o}{S_i/N_i} = 2. \qquad (7.19)$$

Следовательно, в случае ДБП-ПН отношение сигнал/шум на выходе детектора в два раза выше, чем на его входе, т. е. улучшение отношения сигнал/шум при ДБП-ПН равно двум.

Это объясняется следующим образом. Поскольку шум — случайное колебание, он имеет синусоидальную и косинусоидальную составляющие [ф-ла (7.1)], которые в детекторе умножаются на $\cos\omega_c t$. Это приводит к сдвигу спектра синусоидальной составляющей на частоту $2\omega_c$, после чего она полностью отфильтровывается фильтром нижних частот. Таким образом, половина шумовой мощности отфильтровывается и отношение сигнал/шум улучшается в два раза.

СИСТЕМА ОБП-ПН

Приемник сигналов ОБП-ПН подобен приемнику сигналов ДБП-ПН, за исключением входного полосового фильтра (см. рис. 3.20). В случае ДБП-ПН фильтр должен пропускать все составляющие в пределах $\omega_c \pm \omega_m$. С другой стороны, в системе ОБП-ПН входной сигнал имеет только одну боковую полосу и, следовательно, полоса пропускания фильтра равна половине полосы пропускания фильтра в случае ДБП-ПН.

Спектры сигналов на входе и выходе детектора приемника сигналов ОБП-ПН представлены на рис. 7.7. На рис. 7.7*а* показан спектр $F(\omega)$ сообщения $f(t)$. Когда это сообщение поступает на

вход модулятора (рис. 3.18), на выходе получается сигнал ОБП-ПН, спектр которого показан на рис. 7.7б [см. ф-лу (3.27)]. Этот сигнал и есть сигнал $f_i(t)$ на входе детектора. Спектр сигнала на выходе детектора показан на рис. 7.7в [см. ф-лу (3.31)].

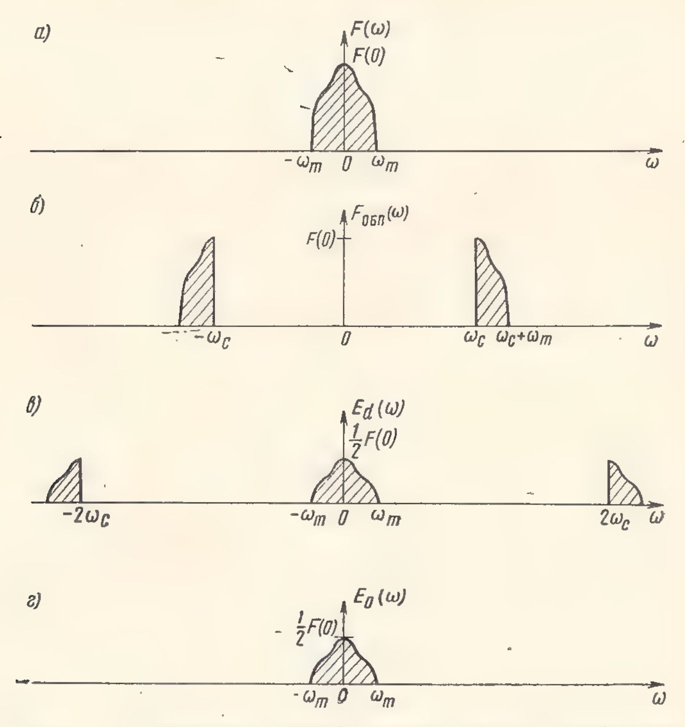


Рис. 7.7. К пояснению преобразования спектра сообщения $f(t)$ в системе передачи ОБП-ПН:
а) спектр сообщения $f(t)$; *б*) спектр сигнала ОБП-ПН; *в*) спектр сигнала на выходе перемножителя; *г*) спектр сигнала на выходе *ФНЧ* детектора

Этот сигнал пропускается через фильтр нижних частот, который отфильтровывает составляющие вблизи $\pm 2\omega_c$. Спектр на выходе фильтра нижних частот показан на рис. 7.7*г*.

Мощность сигнала равна площади под кривой спектра плотности мощности, деленной на 2π. Из рис. 7.7*а* и *б* видно, что площади под кривыми спектров сигналов $f(t)$ и $f_{ОБП}(t)$ одинаковы. Следовательно, одинаковы и мощности однополосного модулированного сигнала (рис. 7.7б) и исходного сообщения $f(t)$. Таким образом, мощность сигнала на входе детектора

$$S_i = \overline{f^2(t)}. \qquad (7.20)$$

Сигнал на выходе детектора, как видно из ф-лы (3.31), равен $\frac{1}{2}f(t)$, следовательно, его мощность

$$S_o=\overline{\left[\frac{1}{2}f(t)\right]^2}=\frac{1}{4}\overline{f^2(t)}=\frac{1}{4}S_i.$$

Таким образом,

$$S_o/S_i=1/4. \tag{7.21}$$

Для определения значений мощностей шума N_i и N_o заметим, что синхронные детекторы сигналов ДБП и ОБП одинаковы. В каждом случае детектирование осуществляется посредством умножения входного сигнала на $\cos\omega_c t$ и фильтрации для устранения спектров вблизи частот $\pm 2\omega_c$. Шумовое колебание на входе детектора является узкополосным. Очевидно, отношение N_o/N_i в приемниках сигналов ДБП и ОБП должно быть одинаково. Из ф-лы (7.18) имеем

$$N_o/N_i=1/4. \tag{7.22}$$

Таким образом, для системы ОБП-ПН

$$\frac{S_o/N_o}{S_i/N_i}=1. \tag{7.23}$$

Отсюда заключаем, что отношения сигнал/шум (по мощности) на входе и выходе детектора одинаковы. Следовательно, отношение сигнал/шум в этом случае не улучшается. На первый взгляд это создает впечатление, что система ДБП превосходит систему ОБП, так как улучшение отношения сигнал/шум в ДБП в два раза больше, чем в ОБП. Однако более внимательное рассмотрение показывает, что такое заключение неверно. Сигнал ОБП требует для передачи только половину полосы частот сигнала ДБП. Следовательно, мощность входного шума N_i в системе ДБП в два раза[1]) больше, чем в системе ОБП. Хотя при ДБП отношение сигнал/шум улучшается в два раза, это улучшение при детектировании аннулируется бо́льшим входным шумом. Легко видеть, что при заданной мощности сигнала на входе отношения сигнал/шум на выходе систем ДБП и ОБП одинаковы. Следовательно, ОБП и ДБП обладают одинаковой помехоустойчивостью.

---

[1]) Если спектр шума симметричен относительно $\omega=\omega_c$ (как в случае белого шума), то мощность входного шума в ДБП точно в два раза больше, чем в ОБП. Для несимметричного спектра шума этот множитель может быть больше или меньше двух в зависимости от того, какая боковая полоса подавляется (см. задачу 7.5). Однако обычно ширина спектра модулированного сигнала столь мала, что без значительной ошибки можно считать, что спектральная плотность помехи одинакова в пределах всей полосы частот сигнала.

СИСТЕМА АМ С НЕСУЩИМ КОЛЕБАНИЕМ (ДЕТЕКТОР ОГИБАЮЩЕЙ)

В случае амплитудной модуляции с большим уровнем несущей сигнал на входе детектора записывается как

$$f_i(t)=[A+f(t)]\cos\omega_c t+n_i(t). \tag{7.24}$$

Полезный сигнал в этом случае равен $[A+f(t)]\cos\omega_c t$. Очевидно, мощности сигнала и шума на входе детектора

$$S_i=A^2/2+\overline{f^2(t)}/2,\quad N_i=\overline{n_i^2(t)}. \tag{7.25}$$

Для вычисления мощностей S_o и N_o найдем огибающую сигнала $f_i(t)$. Представив узкополосный шум (7.24) по ф-ле (7.1), получим

$$f_i(t)=[A+f(t)]\cos\omega_c t+n_c(t)\cos\omega_c t+n_s(t)\sin\omega_c t=$$
$$=[A+f(t)+n_c(t)]\cos\omega_c t+n_s(t)\sin\omega_c t=$$
$$=E(t)\cos[\omega_c t+\psi(t)], \tag{7.26}$$

где

$$E(t)=\sqrt{[A+f(t)+n_c(t)]^2+n_s^2(t)}, \tag{7.27}$$

$$\psi(t)=-\operatorname{arc\,tg}\frac{n_s(t)}{A+f(t)+n_c(t)}. \tag{7.28}$$

Очевидно, $E(t)$ есть огибающая сигнала $f_i(t)$, а $\psi(t)$ — фазовый угол. Выходное напряжение детектора огибающей равно $E(t)$. Рассмотрим два случая: малый шум, когда $A+f(t)\gg n_i(t)$, и большой шум, когда $n_i(t)\gg A+f(t)$.

Случай малого шума. Используем векторную диаграмму сигнала $f_i(t)$, определяемого ф-лой (7.26) (см. рис. 7.8). Если $A+$

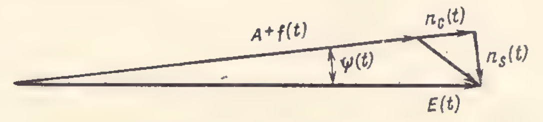


Рис. 7.8. Векторная диаграмма суммы сигнала АМ и узкополосного шума при большом отношении сигнал/шум

$+f(t)\gg n_i(t)$, то $A+f(t)\gg n_c(t)$ и $A+f(t)\gg n_s(t)$. Суммарный вектор $E(t)$ в этом случае

$$E(t)\approx A+f(t)+n_c(t)$$

и

$$\psi(t)\approx 0.$$

Получим этот же результат аналитически. Если $A+f(t) \gg n_c(t)$ и $n_s(t)$, то ф-ла (7.27) приближенно выражается в виде

$$E(t) \approx \sqrt{[A+f(t)]^2+2[A+f(t)]\,n_c(t)} =$$

$$=[A+f(t)]\left[1+\frac{2\,n_c(t)}{A+f(t)}\right]^{\frac{1}{2}} \approx [A+f(t)]\left[1+\frac{n_c(t)}{A+f(t)}\right] =$$

$$=A+f(t)+n_c(t).$$

Из этого выражения вытекает, что полезный сигнал на выходе детектора равен $f(t)$, а шум — $n_c(t)$. Следовательно,

$$S_o=\overline{f^2(t)}, \quad N_o=\overline{n_c^2(t)}=N_i. \tag{7.29}$$

Используя ф-лы (7.25) и (7.29), получаем

$$\frac{S_o/N_o}{S_i/N_i}=\frac{2\,\overline{f^2(t)}}{A^2+\overline{f^2(t)}}. \tag{7.30}$$

Улучшение отношения сигнал/шум растет по мере уменьшения A. Но для детектора огибающей A не может быть меньше $|f(t)|_{\text{макс}}$. Легко видеть, что отношение мощностей сигнала и шума на выходе приемника АМ достигает наибольшего значения при наивысшем допустимом коэффициенте модуляции (100%).

В частном случае, когда $f(t)$ — гармоническая функция, амплитуда $f(t)$ при 100-процентной модуляции равна A. Следовательно,

$$\overline{f^2(t)}=\frac{A^2}{2}, \quad \frac{S_o/N_o}{S_i/N_i}=\frac{2}{3}.$$

Итак, наибольшее улучшение отношения сигнал/шум по мощности, которое можно достичь в этом случае, равно 2/3.

Если для детектирования сигнала АМ с большим уровнем несущей используется синхронный детектор, то получается такой же результат, как и для детектора огибающей при малом шуме. Это объясняется тем, что S_i и N_i в обоих случаях одинаковы:

$$S_i/N_i=[A^2+\overline{f^2(t)}]/[2\,\overline{n_i^2(t)}].$$

При синхронном детектировании входной сигнал $f_i(t)$ умножается на $\cos\omega_c t$. Следовательно, сигнал на выходе перемножителя

$$u_d(t)=f_i(t)\cos\omega_c t.$$

Подставив в это выражение ф-лу (7.26) и устранив слагаемые, спектры которых лежат вблизи частот $\pm 2\omega_c$, получим выходное колебание

$$u_0(t)=\frac{1}{2}[A+f(t)+n_c(t)]. \tag{7.31}$$

В этом выходном колебании полезный сигнал равен $\frac{1}{2} f(t)$, а шум — $\frac{1}{2} n_c(t)$. Следовательно,

$$S_o = \frac{1}{4} \overline{f^2(t)}, \quad N_o = \frac{1}{4} \overline{n_c^2(t)} = \frac{1}{4} \overline{n_i^2(t)}.$$

Таким образом,

$$\frac{S_o}{N_o} = \frac{\overline{f^2(t)}}{\overline{n_i^2(t)}}$$

и

$$\frac{S_o/N_o}{S_i/N_i} = \frac{2\overline{f^2(t)}}{A^2 + \overline{f^2(t)}}, \qquad (7.32)$$

что совпадает с ф-лой (7.30).

Отсюда видно, что для системы АМ в случае малого шума помехоустойчивости детектора огибающей и синхронного детектора одинаковы. Заметим, что при выводе ф-лы (7.32) не делались какие-либо предположения об относительных значениях сигнала и шума. Следовательно, ф-ла (7.32) для синхронного детектирования справедлива при любой мощности шума.

Случай большого шума. Рассмотрим помехоустойчивость детектирования АМ с помощью детектора огибающей, когда $n_i(t) \gg [A + f(t)]\cos\omega_c t$, т. е. когда $n_c(t)$ и $n_s(t) \gg [A + f(t)]$. В этом случае ф-ла (7.27) принимает вид

$$E(t) \approx \sqrt{n_c^2(t) + n_s^2(t) + 2n_c(t)[A + f(t)]} =$$

$$= R(t)\sqrt{1 + \frac{2[A + f(t)]}{R(t)}\cos\theta(t)}, \qquad (7.33)$$

где $R(t)$ и $\theta(t)$ — огибающая и фаза колебания $n_i(t)$ [ф-лы (7.10), (7.11)].

Поскольку $R(t) \gg [A + f(t)]$, ф-лу (7.33) можно приближенно выразить как

$$E(t) \approx R(t)\left[1 + \frac{A + f(t)}{R(t)}\cos\theta(t)\right] = R(t) + [A + f(t)]\cos\theta(t). \qquad (7.34)$$

Из ф-лы (7.34) следует, что выход детектора не содержит неискаженного сигнала $f(t)$. Колебание $f(t)\cos\theta(t)$ представляет собой сигнал $f(t)$, умноженный на шумовое колебание $\cos\theta(t)$, и из него нельзя выделить $f(t)$. Итак, неискаженный полезный сигнал $f(t)$ в выходном колебании детектора не содержится.

Из изложенного видно, что в случае большого шума детектор огибающей полностью подавляет сигнал. Это учитывается так называемым пороговым эффектом в детекторе огибающей. Порогом называется значение отношения сигнал/шум на входе детектора, ниже которого выходное отношение сигнал/шум убывает значительно быстрее, чем входное. Подавление сигнала шумом в детекторе огибающей становится заметным, когда отношение мощно-

стей несущей и шума на входе детектора приближается к единице.

Следует подчеркнуть, что пороговый эффект — это свойство детектора огибающей. Он отсутствует в синхронном детекторе. Улучшение отношения сигнал/шум, определяемое ф-лой (7.32), сохраняется при любом шуме[1]).

Итак, если сигнал АМ передается при малом уровне шума, то помехоустойчивость детектора огибающей почти равна помехоустойчивости синхронного детектора. Но при большом шуме в детекторе огибающей возникает пороговый эффект и он становится хуже синхронного детектора.

7.4. Помехоустойчивость систем связи с угловой модуляцией

ЧАСТОТНАЯ МОДУЛЯЦИЯ

Структурная схема системы связи с частотной модуляцией показана на рис. 7.9. Входной фильтр в приемнике отфильтровывает шумовые составляющие, частоты которых лежат вне полосы ($\omega_c \pm \pm \Delta\omega$), занимаемой спектром полезного сигнала. Если девиация частоты равна $\Delta\omega$, то, очевидно, полоса пропускания этого фильтра согласно ф-ле (4.27) равна $2\Delta\omega$. Колебание $u_d(t)$ на входе *ФНЧ* детектора состоит из полезного сигнала и шума с шириной

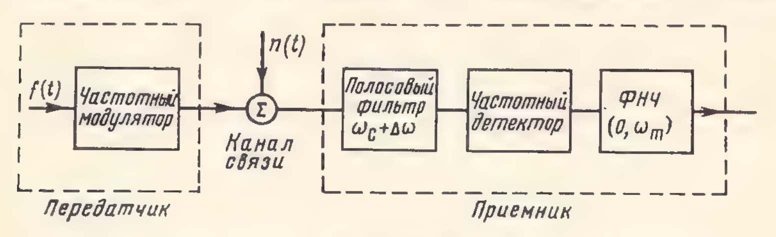


Рис. 7.9. Структурная схема системы связи с частотной модуляцией

спектра $\Delta\omega$. Поскольку ширина спектра полезного сигнала равна ω_m, можно отфильтровать шумы вне этой полосы фильтром нижних частот с частотой среза ω_m (рис. 7.9).

При вычислении мощностей сигнала и шума на выходе приемника будем предполагать, что одна из них может быть найдена независимо от другой. Таким образом, при определении мощности сигнала будем предполагать, что шум в канале равен нулю, а при вычислении мощности шума, что сообщение $f(t)$ равно нулю. Справедливость такого подхода показана ниже.

---

[1] В системах ДБП-ПН и ОБП-ПН (которые используют синхронные детекторы) пороговый эффект также отсутствует.

Сначала рассмотрим сигнал в отсутствие шума. Сигнал частотной модуляции

$$f_i(t) = A\cos\left[\omega_c t + k_f \int f(t)\,dt\right].$$

В § 4.7 отмечалось, что мощность сигнала как при модуляции, так и в ее отсутствие одинакова и равна $S_i = A^2/2$. Выходное напряжение детектора пропорционально мгновенной частоте сигнала ω_i. Если коэффициент пропорциональности обозначить через α, то выходное напряжение сигнала будет выражаться как

$$s_o(t) = \alpha\,\omega_i = \alpha \frac{d}{dt}\left[\omega_c t + k_f \int f(t)\,dt\right] = \alpha\,\omega_c + \alpha\,k_f\,f(t).$$

Полезный сигнал равен $\alpha k_f f(t)$ и

$$S_o = \alpha^2 k_f^2 \overline{f^2(t)}. \tag{7.35}$$

Для того чтобы определить значения N_i и N_o заметим, что ширина спектра сигнала на входе детектора равна $2\Delta\omega$, где $\Delta\omega$ — наибольшее отклонение частоты [см. ф-лу (4.27)].
Таким образом,

$$N_i = \frac{1}{\pi}\int\limits_{\omega_c - \Delta\omega}^{\omega_c + \Delta\omega} S_n(\omega)\,d\omega, \tag{7.36}$$

где $S_n(\omega)$ — спектр плотности мощности колебания $n_i(t)$.

Если шум — белый и имеет плотность мощности $N/2$, то

$$N_i = \frac{1}{\pi}\int\limits_{\omega_c - \Delta\omega}^{\omega_c + \Delta\omega} \frac{N}{2}\,d\omega = 2N\Delta f, \quad (\Delta\omega = 2\pi\Delta f). \tag{7.37}$$

При определении N_o будем предполагать, что сообщение $f(t)$ равно нулю. Колебание на входе детектора есть сумма несущей и шума:

$$f_i(t) = A\cos\omega_c t + n_i(t) = [A + n_c(t)]\cos\omega_c t + n_s(t)\sin\omega_c t =$$
$$= E(t)\cos[\omega_c t + \psi(t)], \tag{7.38}$$

где

$$E(t) = \sqrt{[A + n_c(t)]^2 + n_s^2(t)}, \tag{7.39}$$

$$\psi(t) = -\operatorname{arc\,tg}\frac{n_s(t)}{A + n_c(t)}. \tag{7.40}$$

Будем предполагать, что шум мал по сравнению с несущей, $A \gg n_c(t)$, и $A \gg n_s(t)$. Следовательно,

$$\psi(t) = -\operatorname{arc\,tg}(n_s(t)/A) \approx -n_s(t)/A. \tag{7.41}$$

Мгновенная частота входного колебания на основании ф-л (7.38) и (7.41)

$$\omega_i = \frac{d}{dt}[\omega_c t + \psi(t)] = \omega_c + \dot{\psi}(t) = \omega_c - \frac{\dot{n}_s(t)}{A}. \qquad (7.42)$$

Напряжение на выходе детектора (до низкочастотной фильтрации)

$$f_d(t) = \alpha\,\omega_i = \alpha\,[\omega_c' - \dot{n}_s(t)/A].$$

Очевидно, составляющая выходного шума

$$n_o(t) = -\alpha\,\dot{n}_s(t)/A. \qquad (7.43)$$

Если колебание $n_s(t)$ имеет спектр плотности мощности $S_{n_s}(\omega)$, то его производная $\dot{n}_s(t)$ имеет спектр $S_{\dot{n}_s}(\omega)$, определяемый по ф-ле (2.39):

$$S_{\dot{n}_s}(\omega) = \omega^2 S_{n_s}(\omega),$$

и спектр плотности мощности выходного шума:

$$S_{n_o}(\omega) = \frac{\alpha^2}{A^2} S_{\dot{n}_s}(\omega) = \frac{\alpha^2\omega^2}{A^2} S_{n_s}(\omega). \qquad (7.44)$$

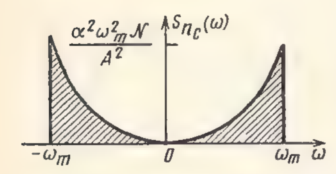


Рис. 7.10. Спектр плотности мощности шума на выходе *ФНЧ* детектора ЧМ

Выходное напряжение детектора пропускается через фильтр нижних частот с частотой среза ω_m, чтобы устранить избыточный шум за пределами спектра сообщения.

После подстановки ф-л (7.6) и (7.7) в выражение (7.44) с учетом фильтрации выходного напряжения фильтром нижних частот получим

$$S_{n_o}(\omega) = \begin{cases} \frac{\alpha^2\omega^2}{A^2}[S_n(\omega+\omega_c) + S_n(\omega-\omega_c)] & \text{при} \quad |\omega| < \omega_m, \\ 0 & \text{при} \quad |\omega| > \omega_m. \end{cases} \qquad (7.45)$$

Если шум является белым, то $S_n(\omega) = N/2$ и

$$S_{n_o}(\omega) = \begin{cases} \frac{\alpha^2\omega^2 N}{A^2} & \text{при} \quad |\omega| < \omega_m, \\ 0 & \text{при} \quad |\omega| > \omega_m. \end{cases} \qquad (7.46)$$

Этот спектр имеет форму параболы (рис. 7.10). Мощность шума на выходе *ФНЧ* детектора

$$N_o = \frac{\alpha^2 N}{\pi A^2}\int_0^{\omega_m} \omega^2\,d\omega = \frac{2}{3}\left(\frac{\alpha}{A}\right)^2 \frac{N\omega_m^3}{2\pi}. \qquad (7.47)$$

Воспользовавшись ф-лами (7.35), (7.37) и (7.47), получим

$$\frac{S_o/N_o}{S_i/N_i}=\frac{6\,k_f^2\overline{f^2(t)}\,\Delta\omega}{\omega_m^3}. \qquad (7.48)$$

Заметим, что отношение мощностей сигнала и шума на выходе детектора согласно ф-лам (7.35) и (7.47)

$$\frac{S_o}{N_o}=\frac{3}{2}\cdot\frac{2\pi A^2}{N}\,\frac{k_f^2\overline{f^2(t)}}{\omega_m^2}. \qquad (7.49)$$

Девиация частоты, а следовательно, и ширина спектра сигнала ЧМ пропорциональны k_f. Поэтому, как вытекает из ф-лы (7.49), отношение мощностей сигнала и шума на выходе частотного детектора пропорционально квадрату ширины спектра сигнала.

Покажем теперь, что независимое определение мощностей сигнала и шума на выходе частотного детектора при слабом входном шуме действительно справедливо.

Шум на входе детектора имеет полосовой спектр, ширина которого приближенно равна $2\Delta\omega$. С другой стороны, несущая частота изменяется пропорционально модулирующему сообщению $f(t)$, спектр которого ограничен частотой ω_m. В случае широкополосной ЧМ $\Delta\omega\gg\omega_m$. Следовательно, шум изменяется значительно быстрее, чем несущая частота. Поэтому разумно допустить, что несущая частота постоянна на интервале $1/(2f_m)$, *сек*. Это равноценно представлению $f(t)$ в виде ступенчатой функции, как показано на рис. 4.3. Таким образом, сигнал имеет постоянную частоту между любыми двумя последовательными отсчетами. Рассмотрим один такой интервал. Частота сигнала на данном интервале постоянна. Это можно интерпретировать так, как если бы модулирующее сообщение $f(t)=0$, а несущая частота $\omega_k=\omega_c+k_f f(t_k)$. Полосовой шум $n_i(t)$, если считать ω_k центральной частотой, выражается как

$$n_i(t)=n_c(t)\cos\omega_k t+n_s(t)\sin\omega_k t.$$

Используя ф-лу (7.43), получим

$$S_{n_o}(\omega)=\frac{\alpha^2\omega^2}{A^2}S_{n_s}(\omega),$$

где

$$S_{n_s}(\omega)=\begin{cases}S_n(\omega-\omega_k)+S_n(\omega+\omega_k) & \text{при}\quad |\omega|<\Delta\omega+2\omega_m+\kappa_f f(t_k),\\ 0 & \text{при}\quad |\omega|>\Delta\omega+2\omega_m+\kappa_f f(t_k).\end{cases}$$

Графики спектров $S_n(\omega)$ и $S_{n_s}(\omega)$ в случае белого шума показаны на рис. 7.11.

Заметим, что ширина спектра передаваемого сигнала равна $2(\Delta\omega+2\omega_m)$. Для белого шума $S_n(\omega)=N/2$. Фильтр нижних частот детектора подавляет все шумы на частотах $|\omega|>\omega_m$. Если $S_n(\omega)=N/2$, то, очевидно, плотность мощности выходного шума $S_{n_o}(\omega)$ определяется как

$$S_{n_o}(\omega)=\begin{cases}\dfrac{\alpha^2\omega^2N}{A^2} & \text{при}\quad |\omega|<\omega_m,\\ 0 & \text{при}\quad |\omega|>\omega_m.\end{cases}$$

Этот результат совпадает с ф-лой (7.46), поэтому и мощность шума в каждом интервале между отсчетами одинакова и определяется по ф-ле (7.47). Видим, что мощность шума не зависит от времени.

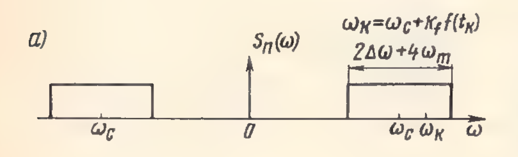
Следует заметить, что приведенное здесь доказательство справедливо только при $\Delta\omega \gg \omega_m$ (т. е. для широкополосной ЧМ) и белом шуме.

Интересно сравнить отношения сигнал/шум на выходе детекторов АМ и ЧМ. Если сообщение $f(t)$ передается посредством АМ, то мощности сигнала и шума на выходе детектора огибающей определяются по ф-ле (7.29). При амплитудной модуляции ширина спектра сигнала равна $2f_m$. Следовательно, для белого шума с плотностью мощности $N/2$

$$N_o = N_i = 2Nf_m \quad \text{и} \quad (S_o/N_o)_{\text{АМ}} = \overline{f^2(t)}/(2Nf_m).$$

Подставив этот результат в ф-лу (7.49), получим

$$\frac{(S_o/N_o)_{\text{ЧМ}}}{(S_o/N_o)_{\text{АМ}}} = 3\left(\frac{Ak_f}{\omega_m}\right)^2. \qquad (7.50)$$



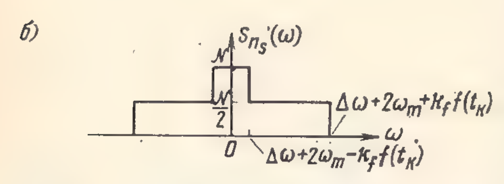


Рис. 7.11. К пояснению независимого определения мощностей сигнала и шума на выходе частотного детектора:
а) спектр входного шума; б) спектр выходного шума

В ф-ле (7.50) A — амплитуда несущей сигнала ЧМ. При сравнении АМ и ЧМ будем предполагать, что амплитуда несущей сигнала АМ также равна A. Чтобы получить количественный результат, рассмотрим частный случай, когда модулирующая функция $f(t)$ представляет собой гармоническое колебание, а коэффициент модуляции $m=1$. Тогда амплитуда сообщения $f(t)$ равна A. Девиация частоты сигнала ЧМ, когда $f(t)$ — гармоническое колебание с амплитудой A, находится по ф-ле (4.20):

$$\Delta\omega = Ak_f$$

и индекс модуляции

$$m_f = \Delta\omega/\omega_m = Ak_f/\omega_m,$$

где ω_m — частота модулирующей функции $f(t)$. Подстановка в ф-лу (7.50) дает

$$\frac{(S_o/N_o)_{\text{ЧМ}}}{(S_o/N_o)_{\text{АМ}}} = 3m_f^2. \qquad (7.51а)$$

Если рассматривается отношение эффективных значений напряжения сигнала и помехи, а не мощностей, то

$$\frac{(s_o/\sigma_o)_{\text{ЧМ}}}{(s_o/\sigma_o)_{\text{АМ}}} = \sqrt{3}\,m_f. \qquad (7.51б)$$

Из ф-л (7.51) видно, что с увеличением индекса модуляции m_f можно получить при ЧМ значительно большее отношение сигнал/шум, чем при АМ. Заметим, однако, что увеличение m_f приводит к росту требуемой для передачи полосы частот. Например, если $m_f=5$, то $(S_o/N_o)_{\text{ЧМ}}$ в 75 раз выше, чем при АМ, но требуемая для передачи полоса частот увеличивается почти в 8 раз. При АМ полоса частот сигнала равна 2 ω_m. При ЧМ с индексом модуляции $m_f=5$ полоса частот сигнала равна 3,3 $\Delta\omega$ (рис. 4.5), но $\Delta\omega=m_f\omega_m$. Следовательно, полоса частот для передачи сигнала ЧМ равна $3{,}3\times5\ \omega_m\approx16\ \omega_m$.

При больших значениях $m_f (m_f>10)$ ширина спектра сигнала ЧМ $B=2\Delta f$ и $m_f=\dfrac{\Delta f}{f_m}=\dfrac{B}{2f_m}$.

Поэтому

$$\frac{(S_o/N_o)_{\text{ЧМ}}}{(S_o/N_o)_{\text{АМ}}}=\frac{3}{4}\left(\frac{B}{f_m}\right)^2. \qquad (7.52)$$

Из этой формулы имеем:

$$\left(\frac{S_o}{N_o}\right)_{\text{ЧМ}}=\frac{3}{4}\left(\frac{B}{f_m}\right)^2\left(\frac{S_o}{N_o}\right)_{\text{АМ}}.$$

Для того чтобы увеличить отношение S_o/N_o в системе ЧМ в четыре раза (6 *дб*), необходимо расширить полосу частот B в два раза. Таким образом, отношение сигнал/шум по мощности улучшается на 6 *дб* при каждом двукратном увеличении занимаемой сигналом полосы частот. Напомним, однако, что эти результаты получены в предположении малого шума.

Перепишем ф-лу (7.52), введя в нее отношение сигнал/шум по напряжению:

$$\frac{(s_o/\sigma_o)_{\text{ЧМ}}}{(s_o/\sigma_o)_{\text{АМ}}}=\frac{\sqrt{3}}{2}\,\frac{B}{f_m}. \qquad (7.53)$$

Таким образом, отношение эффективных значений напряжения сигнала и шума при ЧМ прямо пропорционально полосе частот сигнала.

Свойство обмена полосы частот на отношение сигнал/шум верно для всех систем связи. Преобразование сигнала, в результате которого он занимает более широкую полосу частот, делает его более невосприимчивым к шумам. Теоретическая основа этого свойства раскрывается в гл. 8.

Следует заметить, что из двух рассмотренных систем только ЧМ допускает обмен отношения сигнал/шум на полосу частот. Для систем АМ, где полоса частот передаваемого сигнала постоянна, такой обмен невозможен.

Из ф-лы (7.51б) следует, что отношение сигнал/шум при ЧМ больше, чем при АМ, если выполняется условие $m_f>1/\sqrt{3}\approx0{,}6$. Интересно отметить, что такое значение m_f является примерно

граничным между узкополосной и широкополосной ЧМ. Следовательно, узкополосная ЧМ не обеспечивает заметного улучшения помехоустойчивости по сравнению с АМ. Этого нужно было ожидать, поскольку узкополосная ЧМ и АМ занимают одинаковую полосу частот.

Отношение сигнал/шум при ЧМ не может увеличиваться до бесконечности при расширении полосы частот сигнала. По мере расширения полосы частот увеличивается также входной шум и, очевидно, при некотором значении полосы частот мощности несущей и входного шума достигают одного порядка. Полученные ранее результаты, основанные на предположении малого входного шума, оказываются неверными. При достаточно большом входном шуме наблюдается явление порога, о котором упоминалось выше.

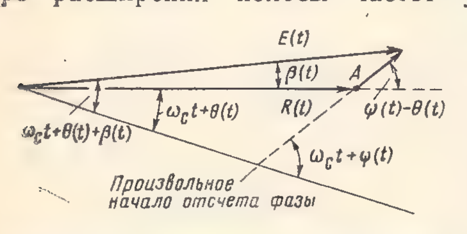


Рис. 7.12. Векторная диаграмма суммы сигнала ЧМ и узкополосного шума при малом отношении сигнал/шум

Пороговый эффект при ЧМ. Пороговый эффект при ЧМ является более заметным и значительным, чем при АМ. Для его анализа выразим колебание на входе детектора в виде

$$f_i(t) = A\cos[\omega_c t + \psi(t)] + n_i(t),$$

где

$$\psi(t) = k_f \int f(t)\, dt.$$

Представив $n_i(t)$ по ф-ле (7.9), получим

$$f_i(t) = A\cos[\omega_c t + \psi(t)] + R(t)\cos[\omega_c t + \theta(t)]. \quad (7.54)$$

Векторная диаграмма, соответствующая ф-ле (7.54), показана на рис. 7.12. Из нее видно, что

$$f_i(t) = E(t)\cos[\omega_c t + \theta(t) + \beta(t)], \quad (7.55)$$

где

$$\beta(t) = \operatorname{arc\,tg} \frac{A\sin[\psi(t) - \theta(t)]}{R(t) + A\cos[\psi(t) - \theta(t)]}. \quad (7.56)$$

В случае большого шума $R(t) \gg A$ и ф-ла (7.56) принимает вид

$$\beta(t) \approx \operatorname{arc\,tg} \frac{A\sin[\psi(t) - \theta(t)]}{R(t)} \approx \frac{A}{R(t)} \sin[\psi(t) - \theta(t)]. \quad (7.57)$$

Заметим, что напряжение на выходе детектора находится по формуле

$$u_d(t) = \alpha \frac{d}{dt}[\omega_c t + \theta(t) + \beta(t)] = \alpha[\omega_c + \dot{\theta}(t) + \dot{\beta}(t)].$$

Напряжение $\alpha\dot{\theta}(t)$ есть напряжение шума. Сведения о передаваемом сообщении $f(t)$ содержатся в $\dot{\beta}(t)$. Из ф-лы (7.57) видно, что $\dot{\beta}(t)$ включает множитель $1/R(t)$, который является случайным колебанием. Следовательно, колебание $\dot{\beta}(t)$ не содержит какого-либо напряжения, пропорционального сообщению $f(t)$, а значит, выходное напряжение детектора является шумовым. Это и приводит к возникновению порогового эффекта, как показано на рис. 7.13.

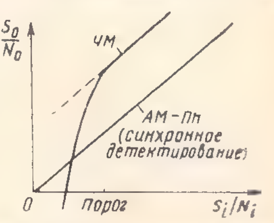


Рис. 7.13. Пороговой эффект при приеме сигналов ЧМ

Улучшение порогового отношения сигнал/шум с помощью предыскажений. Выше отмечалось, что плотность мощности шума на выходе детектора увеличивается пропорционально квадрату частоты (рис. 7.10). Это нежелательно, так как для всех встречающихся на практике сообщений их плотность мощности убывает с ростом частоты. Таким образом, шум оказывается более сильным на тех частотах, где мощность сообщения мала. Поэтому верхние частотные составляющие сообщения подвержены бо́льшим искажениям при передаче. Этот недостаток можно частично устранить с помощью так называемого метода предыскажения и последующей коррекции.

Сущность метода предыскажения состоит в том, что на передающей стороне верхние частотные составляющие сообщения увеличиваются по амплитуде. Такое преобразованное сообщение $f'(t)$ модулирует несущую по частоте. На выходе детектора приемника получаются преобразованное сообщение $f'(t)$ и шум с квадратичным спектром. Для получения неискаженного сообщения это напряжение пропускается через фильтр, который ослабляет верхние частотные составляющие до первоначального уровня (коррекция). В результате получается исходное сообщение $f(t)$ и шум, спектр плотности мощности которого соответственно ослаблен на верхних частотах. Метод предыскажений и последующей коррекции снижает не только шум на тех частотах, где составляющие сигнала малы, но и общий уровень шума на выходе, а также увеличивает выходное отношение сигнал/шум. Это приводит к снижению порога при ЧМ. Определим улучшение отношения сигнал/шум при применении предыскажения сообщения с последующей коррекцией.

Рассмотрим наиболее простой предыскажающий фильтр, показанный на рис. 7.14*а*, и его частотную характеристику. Частоты ω_1 и ω_2 равны $\omega_1=1/(R_1C)$, $\omega_2\approx 1/(R_2C)$, $(R_1\gg R_2)$. Частота ω_1 выбирается равной той частоте, на которой спектральная плотность сообщения $f(t)$ уменьшается на 3 *дб* относительно его значения на низких частотах. Для радиовещательных систем $f_1=\omega_1/2\pi$ выбирается равной 2,1 *кгц*. Частоту ω_2 следует выбирать так, чтобы она значительно превосходила наивысшую частоту со-

общения $f(t)$. В системах радиовещания обычно $f_2=\omega_2/2\pi\geqslant 30$ *кгц*. Соответствующая корректирующая цепь и ее частотная характеристика показаны на рис. 7.14*б*. Поскольку спектр сообщения равен нулю при $\omega>\omega_2$, коррекция на частотах выше ω_2 не требуется. Для такого корректирующего фильтра

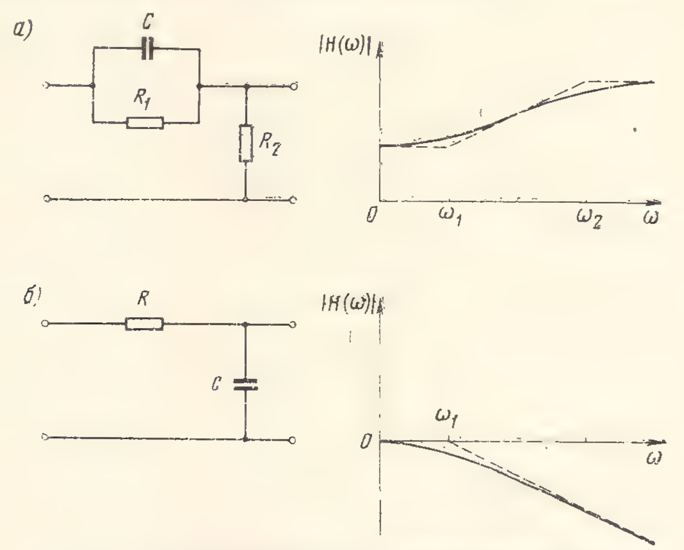


Рис. 7.14. Предыскажающий и корректирующий фильтры: *а*) схема простого предыскажающего фильтра и его амплитудно-частотная характеристика; *б*) схема простого корректирующего фильтра и его амплитудно-частотная характеристика

$$H(\omega)=\frac{1}{1+i\dfrac{\omega}{\omega_1}},\ |H(\omega)|^2=\frac{\omega_1^2}{\omega^2+\omega_1^2}. \tag{7.58}$$

Если $n_o'(t)$ представляет выходной шум (после корректирующего фильтра), а шум в канале передачи является белым, то спектр плотности мощности $S_{n_o'}(\omega)$ с учетом ф-лы (7.46) будет

$$S_{n_o'}(\omega)=\frac{\alpha^2\,\omega^2 N}{A^2}\,|H(\omega)|^2=\frac{\alpha^2\,\omega_1^2\,\omega^2 N}{A^2\left(\omega^2+\omega_1^2\right)}.$$

Мощность выходного шума

$$N_o'=\frac{1}{\pi}\int\limits_0^{\omega_m}S_{n_o'}(\omega)\,d\omega=\frac{\alpha^2\,\omega_1^2 N}{\pi A^2}\int\limits_0^{\omega_m}\frac{\omega^2}{\omega^2+\omega_1^2}\,d\omega=$$

$$=\frac{\alpha^2\,\omega_1^3 N}{\pi A^2}\left[\frac{\omega_m}{\omega_1}-\operatorname{arc\,tg}\frac{\omega_m}{\omega_1}\right]. \tag{7.59}$$

Если определить коэффициент снижения шума как $\rho = N_o/N'_0$, то из ф-л (7.47) и (7.59) получим

$$\rho = \frac{1}{3} \frac{\left(\frac{\omega_m}{\omega_1}\right)^3}{\frac{\omega_m}{\omega_1} - \operatorname{arc\,tg} \frac{\omega_m}{\omega_1}} . \tag{7.60}$$

График коэффициента ρ в зависимости от ω_m/ω_1 приведен на рис. 7.15.

Важно подчеркнуть, что предыскажение не вызывает увеличения мощности передаваемого сигнала, так как при ЧМ мощность передаваемого сигнала равна мощности немодулированной несущей ($A^2/2$). Мощность шума на входе детектора N_i и мощность сигнала на его выходе S_o также не изменяются. Следовательно, ρ в ф-ле (7.60) действительно определяет дополнительное улучшение отношения сигнал/шум при ЧМ.

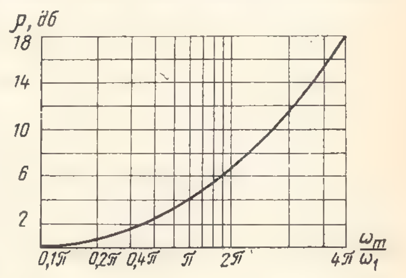


Рис. 7.15. Зависимость коэффициента снижения шума от отношения ω_m/ω_1 в системе с предыскажением

Коэффициент ρ растет по мере увеличения отношения ω_m/ω_1. Это означает, что при уменьшении частоты ω_1 помехоустойчивость возрастает. Но уменьшение частоты ω_1 приводит к расширению передаваемой полосы частот. Это объясняется тем, что по мере уменьшения частоты ω_1 предыскажающий фильтр все более и более приближается к идеальному дифференциатору. Действительно, при $\omega_1 = 0$ предыскажающий фильтр действует как идеальный дифференциатор. Следовательно, исходное сообщение $f(t)$ дифференцируется, прежде чем поступает на вход частотного модулятора. Процесс дифференцирования стремится еще больше увеличить резкие пики большой амплитуды (что обратно эффекту сглаживания при интегрировании). Поскольку постоянная модуляции k_f фиксирована, возросшие максимальные значения сообщения вызывают увеличение девиации частоты $\Delta\omega$. Это приводит к увеличению полосы частот передаваемого сигнала. Следовательно, улучшение отношения сигнал/шум при использовании предыскажения с последующей коррекцией достигается за счет расширения спектра сигнала. Попутно следует заметить, что если $\omega_1 = 0$, то ЧМ с предыскажением и коррекцией эквивалентна фазовой модуляции. Это объясняется тем, что при $\omega_1 = 0$ предыскажающий фильтр действует как дифференциатор. Как указывалось в гл. 4, если сообщение дифференцируется, а затем поступает на частотный модулятор, то получающийся на его выходе сигнал является модулированным по фазе.

Фазовая модуляция может рассматриваться как частный случай ЧМ. Как отмечалось выше, ФМ можно получить, дифференцируя $f(t)$, а затем модулируя несущую по частоте. Если в приемнике используется частотный детектор, выходной сигнал будет df/dt. Неискаженный сигнал можно восстановить, добавив к обычному приемнику ЧМ идеальный интегратор. Сигнал ФМ можно записать как

$$\varphi_{ФМ}(t) = A\cos[\omega_c t + k_p f(t)] = A\cos\left[\omega_c t + k_p \int \frac{df}{dt}\,dt\right],$$

$S_i = A^2/2$; $N_i = 2N\Delta f$ (в случае белого шума).

Сигнал на выходе частотного детектора равен $\alpha k_p(df/dt)$, где α — постоянная. Далее выходной сигнал интегрируется для получения $\alpha k_p f(t)$. Следовательно, $S_o = \alpha^2 k_p^2 \overline{f^2(t)}$. Шум на выходе частотного детектора согласно ф-ле (7.43) равен $[-\alpha \dot{n}_s(t)/A]$. Идеальный интегратор, включенный после частотного детектора, дает на выходе шум $n_o(t)$:

$$n_o(t) = -\alpha n_s(t)/A$$

со спектром

$$S_{n_o}(\omega) = \frac{\alpha^2}{A^2} S_{n_s}(\omega) = \begin{cases} \frac{\alpha^2}{A^2}[S_n(\omega-\omega_c) + S_n(\omega+\omega_c)], & \text{при } |\omega| < \omega_m, \\ 0 & \text{при } |\omega| > \omega_m. \end{cases}$$

Если в канале передачи действует белый шум, то

$$S_n(\omega) = N/2,$$

$$S_{n_o}(\omega) = \begin{cases} \frac{\alpha^2 N}{A^2} & \text{при } |\omega| < \omega_m, \\ 0 & \text{при } |\omega| > \omega_m. \end{cases}$$

Заметим, что, в отличие от ЧМ, выходной шум приемника ФМ имеет равномерный спектр. Следовательно, нет необходимости в предыскажении и последующей коррекции. Действительно, как упоминалось выше, ФМ можно рассматривать как ЧМ с предыскажением и последующей коррекцией.

Мощность выходного шума

$$N_o = \frac{1}{\pi}\int_0^{\omega_m} S_{n_o}(\omega)\,d\omega = \frac{2\alpha^2 N f_m}{A^2}.$$

Следовательно,

$$\frac{S_o}{N_o} = \frac{(k_p A)^2}{2Nf_m}\overline{f^2(t)},$$

$$\frac{S_o/N_o}{S_i/N_i}=\frac{2k_p^2\overline{f^2(t)}\,\Delta f}{f_m}.$$

Наконец,

$$\frac{(S_o/N_o)_{\text{ФМ}}}{(S_o/N_o)_{\text{АМ}}}=(k_p A)^2.$$

Здесь A — амплитуда несущей сигнала ФМ. При сравнении АМ и ФМ допустим, что несущая сигнала АМ имеет амплитуду A и 100-процентную глубину модуляции. В частном случае, когда $f(t)$ — гармоническая функция и глубина модуляции равна 100%, $f(t)=A\cos\omega_m t$:

$$\varphi_{\text{ФМ}}(t)=A\cos(\omega_c t+A k_p\cos\omega_m t),$$

$$\omega_i=\omega_c-k_p A\,\omega_m\sin\omega_m t,$$
$$\Delta\omega=k_p\,\omega_m A.$$

Следовательно,

$$\frac{(S_o/N_o)_{\text{ФМ}}}{(S_o/N_o)_{\text{АМ}}}=\left(\frac{\Delta\omega}{\omega_m}\right)^2.$$

Выраженный через отношения сигнал/шум по напряжению коэффициент улучшения помехоустойчивости

$$\frac{(s_o/\sigma_o)_{\text{ФМ}}}{(s_o/\sigma_o)_{\text{АМ}}}=\frac{\Delta\omega}{\omega_m}.$$

Заметим, что как и при ЧМ, улучшение отношения сигнал/шум по мощности в случае ФМ увеличивается пропорционально квадрату полосы частот сигнала.

7.5. Помехоустойчивость систем связи с импульсной модуляцией

АМПЛИТУДНО-ИМПУЛЬСНАЯ МОДУЛЯЦИЯ

Поведение отношения сигнал/шум при АИМ такое же, как и при АМ-ПН. Улучшение отношения сигнал/шум при АИМ равно единице, точно так же, как при ОБП-ПН. Подобным же образом улучшение отношения сигнал/шум при АИМ/АМ равно двум, как и при ДБП-ПН. Этот результат не является неожиданным, поскольку необходимая полоса частот при АИМ и АИМ/АМ совпадает с ОБП и ДБП соответственно (см. § 5.4).

В системах АИМ сообщение $f(t)$ дискретизируется по времени. Выборочные значения сообщения (отсчеты) $f_s(t)$ затем про-

пускаются через фильтр нижних частот[1]), который восстанавливает исходное непрерывное сообщение. Итак, в действительности, сообщение $f(t)$ передается непосредственно. В процессе передачи сообщение искажается шумом. В приемнике применяется фильтр нижних частот. Очевидно, что отношение сигнал/шум в этом случае остается неизменным. Таким образом,

$$\frac{S_o/N_o}{S_i/N_i} = 1. \tag{7.61}$$

Можно видеть, что $S_o = S_i = \overline{f^2(t)}$, $N_o = N_i = N f_m$ (для белого шума).

Выражение (7.61) справедливо и в случае передачи M сообщений при временно́м уплотнении. Если сигнал АИМ передается как АИМ/АМ, легко можно показать, что улучшение отношения сигнал/шум равно двум (такое же, как при ДБП-ПН).

Рассмотрим систему АИМ с временны́м уплотнением (рис. 7.16), которая уже обсуждалась в § 5.4. Имеется M сообщений $f_1(t)$, $f_2(t)$, ..., $f_M(t)$, причем спектр каждого сообщения ограничен частотой B, *гц*. От каждого из этих сообщений берутся отсчеты, а затем все отсчеты передаются последовательно.

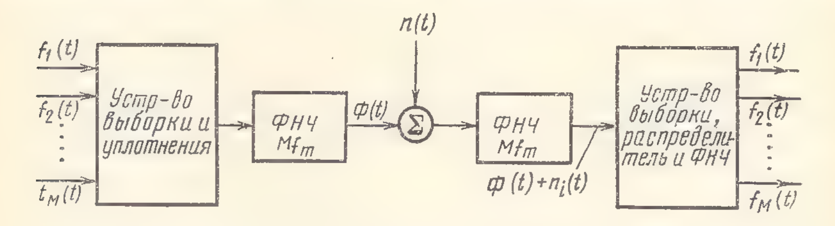


Рис. 7.16. Блок-схема многоканальной системы связи с АИМ и временны́м уплотнением

Затем уплотнённые по времени отсчеты пропускаются через фильтр нижних частот с частотой среза MB, *гц*. Напряжение на выходе этого фильтра представляет непрерывное колебание, значения которого в моменты отсчетов равны значениям отсчетов отдельных сообщений[2]). Обозначим выходное напряжение этого фильтра через $\varphi(t)$. Чтобы найти мощность колебания $\varphi(t)$, используем ф-лу (2.31), которая устанавливает, что среднее квадратическое значение сигнала

---

[1]) В действительности имеется M сообщений, отсчеты которых передаются последовательно с целью временно́го уплотнения. Эта совокупность отсчетов затем пропускается через фильтр нижних частот, имеющий частоту среза Mf_m. Здесь же, для удобства, будем рассматривать случай передачи одного сообщения.

[2]) В действительности значения непрерывного колебания в T раз меньше значений отсчетов, где T — интервал между отсчетами [в данном случае $T = 1/(2MB)$]. Это видно из ф-лы (5.1). Для удобства предположим, что идеальный фильтр имеет коэффициент передачи, равный T, так что выходное напряжение фильтра есть непрерывное колебание, значения которого в моменты отсчетов равны соответствующим отсчетам сообщений. Это предположение никак не влияет на вычисления.

$f(t)$ с ограниченным спектром равно среднему квадратическому значению его отсчетов $\overline{f^2(t)}=\overline{f_k^2}$, где $\overline{f_k^2}$ — среднее квадратическое значение отсчетов. Это соотношение получается посредством суммирования квадратов значений отсчетов (взятых на достаточно большом интервале времени) и делением на число отсчетов. Если колебание $f(t)$ ограничено по спектру частотой B, *гц*, то

$$\overline{f_k^2}=\lim_{\tau\to\infty}\frac{1}{2B\tau}\sum_k f_k^2.$$

Здесь f_k—k-й отсчет колебания $f(t)$; $2B\tau$—общее число отсчетов на интервале τ, *сек* (отсчеты берутся с частотой $2B$ отсчетов в секунду).

Очевидно, среднее квадратическое значение $\varphi(t)$ равно среднему квадратическому значению его отсчетов $\overline{\varphi^2(t)}=\overline{\varphi_k^2}$. Но отсчеты $\varphi(t)$ являются отсчетами M сообщений $f_1(t)$, $f_2(t)$, ..., $f_M(t)$. Для удобства допустим, что все сообщения имеют одинаковые мощности

$$\overline{f_1^2(t)}=\overline{f_2^2(t)}=\ldots=\overline{f_M^2(t)}=\overline{f^2(t)}.$$

Очевидно, среднее значение квадрата каждого из сообщений равно $\overline{f^2(t)}$. Если все эти отсчеты уплотняются, то среднее значение квадрата остается неизменным, так как все отсчеты имеют одно и то же среднее квадратическое значение. Таким образом, $\overline{\varphi^2(t)}=\overline{f^2(t)}$.

Итак, мощность (среднее квадратическое значение) $\varphi(t)$ такая же, как мощность отдельного сообщения. Эта мощность есть мощность входного сигнала $S_i=\overline{f^2(t)}$. Ширина спектра колебания $\varphi(t)$ равна MB, и, следовательно, входной фильтр приемника имеет полосу пропускания MB. Если в канале действует белый шум, имеющий плотность мощности $N/2$, то мощность шума на входе приемника

$$\overline{n_i^2(t)}=2MB\frac{N}{2}=MNB=N_i.$$

Заметим, что среднее квадратическое значение отсчетов шума также равно MNB. Таким образом, $\overline{n_k^2}=MNB$.

В приемнике берутся отсчеты непрерывного колебания $\varphi(t)+n_i(t)$ с частотой $2MB$ отсчетов в секунду. Отсчет в k-й момент равен сумме k—х отсчетов φ_k и $n_k(\varphi_k+n_k)$. Среднее квадратическое значение φ_k равно $\overline{f^2(t)}$, а среднее квадратическое значение n_k равно MNB. Эти отсчеты затем распределяются по M каналам. Среднее квадратическое значение отсчетов в каждом из M каналов остается неизменным. Далее эти отсчеты в каждом канале приемника пропускаются через фильтры нижних частот, что дает M отдельных сообщений. Среднее квадратическое значение каждого сообщения равно $\overline{\varphi_k^2}=\overline{f^2(t)}$, а среднее квадратическое значение шума на выходе каждого канала равно $\overline{n_k^2}=MNf_m$. Следовательно, $S_o=\overline{f^2(t)}$, $N_o=MNB$. Отсюда получаем

$$\frac{S_o/N_o}{S_i/N_i}=1.$$

ФАЗОВО-ИМПУЛЬСНАЯ МОДУЛЯЦИЯ

При ФИМ сообщение $f(t)$ дискретизируется по времени, а затем отсчеты сообщения передаются посредством изменения временно́го положения импульсов. Сдвиг k-го импульса относительно его положения в отсутствие модуляции пропорционален $f(kT)$,

где $f(kT)$ k-й отсчет сообщения $f(t)$ (см. рис. 5.9). Импульсы передаются по каналу с шириной полосы пропускания B. Это приводит к «расплыванию» импульса во времени, как показано в § 2.6 (рис. 2.9). Принимаемый импульс с достаточной степенью

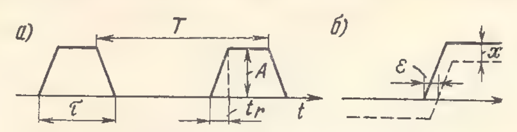


Рис. 7.17. Действие аддитивного шума на сигнал ФИМ:
а) трапецеидальные импульсы сигнала; *б)* изменение временного положения импульса под действием шума

точности можно представить в форме трапеции (рис. 7.17 см. также рис. 2.9). Время нарастания импульса t_r определяется по ф-ле (2.11):

$$t_r \approx 1/B,$$

где B — ширина полосы частот канала, *гц*.

Положение трапецеидального импульса чувствительно к аддитивному шуму канала. Это видно из рис. 7.17*б*. Прибавление постоянного напряжения величины x изменяет положение импульса на ε, где

$$\varepsilon/x = t_r/A. \tag{7.62}$$

Если x изменяется случайным образом, то ε также изменяется случайно. Поскольку значение ε пропорционально x, то эффективное значение ε будет пропорционально эффективному значению x. Из ф-лы (7.62) следует, что

$$\overline{\varepsilon^2}/\overline{x^2} = (t_r/A)^2. \tag{7.63}$$

Таким образом, если канальный шум $n(t)$, имеющий среднее квадратическое значение $\overline{n^2(t)}$, добавляется к сигналу ФИМ, то среднее квадратическое значение ε

$$\overline{\varepsilon^2} = (t_r/A)^2\,\overline{n^2(t)}. \tag{7.64}$$

В отсутствие шума положение k-го импульса изменяется пропорционально значению $f(kT)$. Если обозначить изменение положения k-го импульса при модуляции через y_k, то

$$y_k = k_1 f(kT),$$

где k_1 — коэффициент пропорциональности.

Однако, благодаря канальному шуму, изменение положения импульса равно не y_k, а $y_k+\varepsilon_k$, где ε_k — случайная величина, имеющая среднее квадратическое значение, определяемое по ф-ле (7.64). Пусть

$$\hat{y}_k = y_k + \varepsilon_k = k_1 f(kT) + \varepsilon_k. \tag{7.65}$$

На приемной стороне положения импульсов преобразуются обратно в отсчеты. Значение k-го отсчета в приемнике равно $\hat{y}_k$ [ф-ла (7.65)]. Чтобы получить непрерывное сообщение $f(t)$, отсчеты пропускаются через фильтр нижних частот с частотой среза ω_m (см. рис. 1.35). Напряжение на выходе фильтра определяется по ф-ле (1.40*б*):

$$\hat{y}(t) = \sum_k \hat{y}_k Sa(\omega_m t - k\pi) = \sum_k [k_1 f(kT) + \varepsilon_k] Sa(\omega_m t - k\pi) =$$

$$= \sum_k k_1 f(kT)\, Sa(\omega_m t - k\pi) + \sum_k \varepsilon_k Sa(\omega_m t - k\pi).$$

Каждая из этих сумм представляет собой колебание со спектром, ограниченным частотой ω_m, *рад/сек*. Поскольку $f(kT)$ — отсчеты сообщения $f(t)$, то первая сумма равна $k_1 f(t)$. Вторая сумма представляет собой шум $\varepsilon(t)$. Таким образом,

$$\hat{y}_k(t) = k_1 f(t) + \varepsilon(t).$$

Следовательно, выходной сигнал равен $s_o(t) = k_1 f(t)$, а его средняя мощность $S_o = k_1^2 \overline{f^2(t)}$; выходной шум $n_o(t) = \varepsilon(t)$, средняя мощность которого $N_o = \overline{n_0^2(t)} = \overline{\varepsilon^2(t)}$.

Теперь используем результат примера 2.2 [ф-ла (2.31)], который устанавливает, что среднее квадратическое значение сигнала равно среднему квадратическому значению его отсчетов. Таким образом, $\overline{\varepsilon^2(t)} = \overline{\varepsilon_k^2}$, которое, в свою очередь, определяется по ф-ле (7.64). Следовательно,

$$\overline{\varepsilon^2(t)} = \overline{\varepsilon_k^2} = (t_r/A)^2\, \overline{n^2(t)}.$$

Отсюда получаем

$$N_o = (t_r/A)^2\, \overline{n^2(t)},$$

$$\frac{S_o}{N_o} = \frac{k_1^2}{\overline{n^2(t)}} \left(\frac{A}{t_r}\right)^2 \overline{f^2(t)} = k_1^2 \frac{A^2}{\overline{n^2(t)}}\, \overline{f^2(t)}\, B^2,$$

где B — полоса пропускания канала, *гц*.

Если отношение длительности импульса τ к периоду следования T равно α, то мощность входного сигнала равна (в предположении, что $t_r \ll T$) $S_i = \alpha A^2$, а мощность входного шума определяется как $N_i = \overline{n^2(t)}$. Следовательно,

$$\frac{S_o/N_o}{S_i/N_i} = \frac{k_1^2}{\alpha}\, \overline{f^2(t)}\, B^2. \tag{7.66а}$$

Если рассматривается отношение сигнал/шум по напряжению, а не по мощности, то

$$\frac{s_0/\sigma_0}{s_i/\sigma_i} = k_1 \left(\frac{\overline{f^2(t)}}{\alpha} \right)^{\frac{1}{2}} B. \tag{7.66б}$$

Итак, улучшение отношения сигнал/шум (по мощности) пропорционально квадрату полосы пропускания канала. Такая же зависимость была получена и для широкополосной ЧМ [ф-ла (7.49)]. В обеих системах возможен обмен полосы частот канала на отношение сигнал/шум. Отношение сигнал/шум по напряжению прямо пропорционально полосе частот B канала. Это характерная особенность систем без кодирования.

КОДОВО-ИМПУЛЬСНАЯ МОДУЛЯЦИЯ

Системы модуляции можно разделить на два класса: системы без кодирования и системы с кодированием. В системах без кодирования один символ из пространства сообщений преобразуется в один же символ в пространстве модулированных сигналов. Например, при АМ каждое возможное значение передаваемого сообщения преобразуется в соответствующее значение амплитуды модулированного сигнала. Подобным образом при ФИМ каждому значению сообщения соответствует определенное смещение импульса. Однако в системах с кодированием (таких, как КИМ) каждый символ сообщения преобразуется в несколько символов сигнала. Таллер показал[1]), что только в системах с кодированием достигается наиболее эффективный обмен полосы частот на отношение сигнал/шум, который вообще возможен теоретически[2]).

Системы без кодирования (такие, как ЧМ или ФИМ) по своей природе неспособны к эффективному обмену полосы частот на отношение сигнал/шум. Рассмотрим КИМ с точки зрения обмена полосы частот на отношение сигнал/шум.

Двоичная КИМ. При передаче с помощью КИМ передаваемое сообщение дискретизируется по времени. Каждый отсчет сообщения передается кодом, образованным из группы импульсов, как рассматривалось в § 5.2. В общем случае отсчеты могут иметь любые величины в некотором диапазоне непрерывных значений. Поэтому сообщение подвергается квантованию, в результате которого непрерывные значения отсчетов сообщения заменяются

---

[1]) В. Дж. Таллер. Теоретические ограничения скорости передачи информации. Теория информации и ее приложения. Сборник переводов под ред. А. А. Харкевича. Физматгиз, 1959.

[2]) Можно показать, однако, что при использовании оптимального фазового детектора (системы ФАП) получается помехоустойчивость, близкая к верхней теоретической границе (в предположении, что ширина полосы частот передачи стремится к бесконечности и мощность передаваемого сигнала распределена равномерно по полосе). См., например, A. Viterbi. Principles of coherent communication, McGraw—Hill, New York, 1966.

ближайшими к ним значениями разрешенных уровней, как показано на рис. 5.10*а*. Очевидно, квантование вносит ошибки. Передаваемое сообщение является приближенной копией действительного сообщения, причем неопределенность значений сообщения равна интервалу между разрешенными уровнями. Увеличивая число уровней в заданном интервале значений сообщения, эту неопределенность (или ошибку) можно сделать сколь угодно малой.

На рис. 5.10*а* имеются 16 разрешенных уровней с интервалами между ними 0,1 *в*. Первым рассмотрим случай двоичного кода, когда амплитуда импульсов может принимать только значения 0 и *A*, *в*. Как видно из рис. 5.10, для представления отсчета, который может принимать любое из 16 возможных значений, требуется группа из четырех двоичных импульсов. Это следует из того, что четыре двоичных импульса могут образовать $2\times2\times2\times\times2=16$ различных комбинаций. Подобным образом для представления отсчета, который может принимать любое из 32 значений, требуется группа из пяти двоичных импульсов. Легко видеть, что для представления отсчета, могущего принимать любое из M возможных значений, требуется группа из $\log_2 M$ двоичных импульсов. Заметим, что если бы отсчеты передавались без квантования, то потребовалось бы бесконечно большое число импульсов, приходящихся на один отсчет сообщения, так как в этом случае $M=\infty$.

В приемнике комбинации двоичных импульсов преобразуются обратно в квантованные отсчеты, которые затем пропускаются через фильтр нижних частот для получения непрерывного сообщения $f(t)$ (см. рис. 7.18). Допустим, что отсчеты сообщения $f(t)$

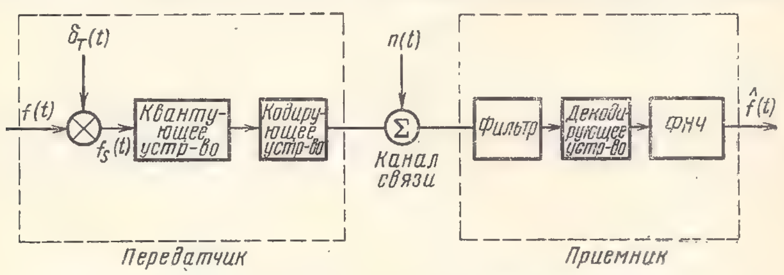


Рис. 7.18 Структурная схема системы связи с КИМ

квантуются на M уровней с интервалами между ними Δu, *в*. Тогда каждый отсчет представляется группой из $\log_2 M$ двоичных импульсов. На рис. 5.10*а* $M=16$, $\Delta u=0{,}1$ *в* и $\log_2 M=4$. Пусть среднее квадратическое значение шума обозначается σ_n^2. Вычислим мощность сигнала на входе декодера приемника.

Входной сигнал состоит из двоичных импульсов. Чтобы найти мощность кодированного сигнала, нужно знать распределение вероятностей значений модулирующего сообщения $f(t)$. Будем пред-

полагать, что значения $f(t)$ распределены равномерно[1]) в заданном диапазоне (от 0 до 1,5 *в* на рис. 5.10*а*). Это означает, что вероятность появления какого-либо значения сообщения та же самая, что и любого другого в заданном диапазоне значений, т. е. равна $1/M$. Отсюда следует, что вероятность появления какой-либо кодовой комбинации равна вероятности появления любой другой кодовой комбинации (для комбинаций на рис. 5.10*б* она равна 1/16). Итак, каждая из кодовых комбинаций появляется одинаково часто. Следовательно, вероятность того, что будет наблюдаться импульс высотой 0 *в*, равна вероятности того, что будет наблюдаться импульс высотой *А, в* (и равна 1/2). В среднем половину времени передачи присутствуют импульсы высотой *А, в*, остальное время импульсы отсутствуют. Отсюда следует, что средняя мощность сигнала определяется как $S_i = A^2/2$. Какой же должна быть высота импульсов A? Очевидно, выбор величины A зависит от шума в канале передачи. Если A значительно больше, чем эффективное значение шума, то обнаружение импульса на фоне шума возможно с очень малой вероятностью ошибки.

В случае двоичной КИМ детектирование сигнала сводится к определению присутствия или отсутствия импульса. Если считать, что $A = K\sigma_n$, то, выбирая достаточно большое значение K, можно сделать вероятность ошибки сколь угодно малой величиной. Этот вопрос рассматривается в гл. 9. Там показано, что при $K=10$ вероятность ошибки имеет порядок 10^{-6}. В настоящем рассмотрении, однако, будем полагать, что K — некоторая неизвестная постоянная. Если вероятности присутствия и отсутствия импульса одинаковы, то $S_i = A^2/2 = K^2\sigma_n^2/2$, $N_i = \sigma_n^2$. Следовательно,

$$S_i/N_i = K^2/2. \tag{7.67}$$

Выбирая достаточно большое значение K, можно сделать вероятность ошибки сколь угодно малой. Следовательно, действие канального шума почти полностью устраняется. Однако остается ошибка в принимаемом сообщении, обусловленная процессом квантования. Неопределенность при квантовании вносится самим передатчиком, и ошибка между действительным и квантованным сообщениями действует как шум. Этот шум называется шумом квантования.

При вычислении шума квантования используем ф-лу (2.31), полученную в примере 2.2. Эта формула устанавливает, что среднее квадратическое значение сигнала с ограниченным спектром равно среднему квадратическому значению его отсчетов взятых через интервал $\leqslant 1/(2f_m)$: $\overline{f^2(t)} = \overline{f_k^2}$. Определим среднее квадратическое значение шума квантования. Каждый отсчет $f(t)$ заменяется ближайшим к нему разрешенным уровнем. Пусть k-й

[1]) Это предположение не обязательно. Можно показать, что полученные здесь результаты справедливы для любого распределения вероятностей значений сообщения при условии, что шаг квантования достаточно мал.

отсчет $f(t)$ есть f_k, а ближайший разрешенный уровень — $\hat{f}_k$. Очевидно, $\hat{f}_k$ — один из M разрешенных уровней, ближайший к f_k. Сообщение $f(t)$ можно выразить через отсчеты f_k по ф-ле (1.140б):

$$f(t) = \sum_k f_k Sa(\omega_m t - k\pi). \tag{7.68а}$$

При декодировании комбинации импульсов в приемнике получается последовательность отсчетов $\hat{f}_k$. Будучи пропущенными через фильтр нижних частот, эти отсчеты дают непрерывную функцию

$$\hat{f}(t) = \sum_k \hat{f}_k Sa(\omega_m t - k\pi). \tag{7.68б}$$

Колебание ошибки $e(t)$ (которое действует подобно шуму):

$$e(t) = f(t) - \hat{f}(t) = \sum_k (f_k - \hat{f}_k) Sa(\omega_m t - k\pi). \tag{7.68в}$$

Если обозначить $e_k = f_k - \hat{f}_k$, то $e(t) = \sum_k e_k S_a(\omega_m t - k\pi)$.

Из результатов, полученных в примере 2.2, следует, что среднее квадратическое значение $e(t)$ равно среднему квадратическому значению его отсчетов e_k: $\overline{e^2(t)} = \overline{e_k^2}$. Заметим, что e_k есть разность между действительным значением отсчета $f(t)$ и его квантованным значением $\hat{f}_k$. Поскольку разрешенные уровни следуют через Δu, в, то e_k лежит в интервале значений $(-\Delta u/2, \Delta u/2)$. Выше предполагалось, что $f(t)$ имеет равномерное распределение в диапазоне $[0, (M-1)\Delta u]$, поэтому распределение e_k также будет равномерным в диапазоне $(-\Delta u/2, \Delta u/2)$. Чтобы вычислить среднее квадратическое значение e_k, разделим весь диапазон $(-\Delta u/2, \Delta u/2)$ на N малых интервалов, каждый шириной $\Delta u/N$. Ошибка e_k будет принимать эти значения, разделенные интервалом $\Delta u/N$, в, одинаково часто. Далее определим среднее значение квадрата всех этих величин[1] в пределе при $N \to \infty$ (или при интервале $\Delta u/N \to 0$).

---

[1]) Читатель, знакомый с теорией вероятностей, может заметить, что $\overline{e_k^2} = \int_{-\infty}^{\infty} e_k^2 p(e_k)\, de_k$, где $p(e_k)$ — функция плотности вероятности величины e_k. В рассматриваемом случае

$$p(e_k) = \begin{cases} 1/\Delta u & -\dfrac{\Delta u}{2} < e_k < \dfrac{\Delta u}{2}, \\ 0 & \text{вне интервала } \left(-\dfrac{\Delta u}{2}, \dfrac{\Delta u}{2}\right). \end{cases}$$

Следовательно, $\overline{e_k^2} = \dfrac{1}{\Delta u} \int_{-\Delta u/2}^{\Delta u/2} e_k^2\, de_k = \dfrac{(\Delta u)^2}{12}$.

Таким образом, в среднем e_k будет одинаково часто принимать значения 0, $\pm\Delta u/N$, $\pm 2\Delta u/N$, ... $\pm r\Delta u/N$, ...$\pm\Delta u/2$. Среднее значение квадрата e_k, очевидно, определяется выражением

$$\overline{e_k^2} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=-N/2}^{N/2} \left(\frac{n\,\Delta u}{N}\right)^2 = \lim_{N\to\infty} \frac{2}{N} \sum_{N=0}^{N/2} \left(\frac{n\,\Delta u}{N}\right)^2 =$$

$$= \lim_{N\to\infty} \frac{2\,(\Delta u)^2}{N^3} \sum_{N=0}^{N/2} n^2 = \lim_{N\to\infty} \frac{2\,(\Delta u)^2}{N^3} \left[\frac{\frac{N}{2}\left(\frac{N}{2}+1\right)(N+1)}{6}\right] = \frac{(\Delta u)^2}{12}.$$

Таким образом, $\overline{e^2(t)} = (\Delta u)^2/12$. Но $\overline{e^2(t)}$ есть среднее квадратическое значение шума квантования на выходе. Обозначим мощность шума квантования через N_q. Следовательно,

$$N_o = N_q = (\Delta u)^2/12. \tag{7.69}$$

Далее определим мощность сигнала на выходе. Выходной сигнал находится как $\hat{f}(t) = \sum_k \hat{f}_k Sa(\omega_m t - k\pi)$.

Используя ф-лу (2.31), имеем $\overline{\hat{f}^2(t)} = \overline{\hat{f}_k^2}$.

Таким образом, среднее квадратическое значение выходного сигнала равно среднему квадратическому значению квантованных отсчетов. Квантованные отсчеты могут принимать M дискретных значений 0, Δu, $2\Delta u$, ..., $(M-1)\Delta u$ с равной вероятностью. Среднее квадратическое значение $\hat{f}_k$ можно записать в виде

$$\overline{\hat{f}_k^2} = \frac{1}{M}\{0^2 + (\Delta u)^2 + (2\,\Delta u)^2 + \ldots + [(M-1)\,\Delta u]^2\} =$$

$$= \frac{(\Delta u)^2}{M} \sum_{k=0}^{M-1} k^2 = \frac{(\Delta u)^2}{M}\,\frac{M(M-1)(2M-1)}{6} = \frac{(\Delta u)^2 (M-1)(2M-1)}{6}.$$

На практике $M \gg 1$, поэтому

$$\overline{\hat{f}_k^2} \approx \frac{M^2}{3}(\Delta u)^2.$$

Следовательно,

$$S_o = \overline{\hat{f}_k^2} = \frac{M^2}{3}(\Delta u)^2. \tag{7.70}$$

Из ф-л (7.67), (7.69) и (7.70) получаем

$$\frac{S_o/N_o}{S_i/N_i} = \frac{4\,M^2}{K^2/2} = \frac{8\,M^2}{K^2}. \tag{7.71}$$

Теперь необходимо связать улучшение отношения сигнал/шум с полосой пропускания системы. При M уровнях квантования требуется $\log_2 M$ импульсов на один отсчет. Поскольку спектр сообщения ограничен частотой f_m, *гц*, то в одну секунду берется $2f_m$ отсчетов и система должна быть в состоянии передавать $2f_m \log_2 M$ импульсов в секунду. В § 5.4 было показано, что система с полосой пропускания B способна передавать $2B$ независимых импульсов в секунду. Следовательно, для передачи $2f_m \log_2 M$ импульсов в секунду требуется полоса пропускания $B = f_m \log_2 M$, откуда $M^2 = 2^{2B/f_m}$.

Таким образом,

$$\frac{S_o/N_o}{S_i/N_i} = \frac{8}{K^2} 2^{2B/f_m}. \tag{7.72}$$

Если рассматривается отношение сигнал/шум по напряжению, то

$$\frac{s_o/\sigma_o}{s_i/\sigma_i} = \frac{\sqrt{8}}{K} 2^{B/f_m}. \tag{7.73}$$

s-ичная КИМ. При двоичной КИМ, как уже указывалось, импульсы могут принимать только два значения. Вместо двоичных могут использоваться s-ичные импульсы, которые могут принимать s значений. Таким образом, при троичной КИМ используются импульсы с тремя амплитудами, например, 0, *в*, A, *в*, и $2A$, *в*. Для представления отсчета, который может принимать любое из M возможных значений, требуется группа из $\log_s M$ импульсов, так как группа из k s-ичных импульсов позволяет образовать s^k различных комбинаций. Следовательно, чтобы образовать M различных комбинаций, необходимо $M = s^k$ и $k = \log_s M$ импульсов на один отсчет. Пусть два ближайших уровня s-ичных импульсов отличаются на $K\sigma_n$. Тогда импульс с r-м уровнем имеет амплитуду $A_r = rk\sigma_n$. Если допустить, что значения $f(t)$ равновероятны в некотором интервале, то вероятность появления в переданном сигнале (использующем s-ичные импульсы) импульса какого-либо уровня будет такая же, как и для любого другого уровня. Следовательно, среднее значение квадрата входного сигнала

$$S_i = \frac{1}{s}[(A_o)^2 + (A_1)^2 + \ldots + (A_{s-1})^2] = \frac{1}{s}\sum_{r=0}^{s-1} A_r^2 =$$

$$= \frac{1}{s}\sum_{r=0}^{s-1} r^2 K^2 \sigma_n^2 = K^2 \sigma_n^2 \frac{(s-1)(2s-1)}{6}. \tag{7.74}$$

В приемнике импульсный сигнал декодируется для получения квантованных отсчетов. Эти отсчеты те же самые, которые получаются при двоичной КИМ. Поэтому мощность выходного сигнала S_o и мощность выходного шума (мощность шума квантования) N_o

равны значениям, найденным для двоичной КИМ [ф-лы (7.70) и (7.69)]. Точно так же $N_i = \sigma_n^2$. Следовательно,

$$\frac{S_o/N_o}{S_i/N_i} = \frac{24\, M^2}{K^2\,(s-1)\,(2s-1)}\,. \tag{7.75}$$

В случае s-ичных импульсов требуется только $\log_s M$ импульсов на один отсчет сообщения, а всего $2f_m \log_s M$ импульсов в секунду. Для их передачи необходима полоса частот $B = f_m \log_s M$. Следовательно, $M^2 = s^{2B/f_m}$ и

$$\frac{S_o/N_o}{S_i/N_i} = \frac{24}{K^2\,(s-1)\,(2s-1)}\, s^{2B/f_m}\,. \tag{7.76}$$

Заметим, что при подстановке в ф-лу (7.76) $s=2$ получим, как и следовало ожидать, ф-лу (7.72).

Отношение сигнал/шум по напряжению

$$\frac{s_o/\sigma_o}{s_i/\sigma_i} = \frac{2\sqrt{6}}{K\sqrt{(s-1)\,(2s-1)}}\, s^{B/f_m}\,. \tag{7.77}$$

Выше рассматривались s-ичные импульсы, амплитуды которых изменялись через $K\sigma_n$ в пределах от 0 до $(s-1)\, K\sigma_n$. Все импульсы имеют положительные значения. Кодированный сигнал, таким образом, имеет постоянную составляющую [в данном случае равную $K\sigma_n\,(s-1)/2$]. Эга постоянная составляющая не несет информации и может не передаваться. Следовательно, можно вычесть эту составляющую и уменьшить мощность сигнала на величину $[K\sigma_n(s-1)/2]^2$. Тогда мощность сигнала

$$S_i' = S_i - \left[\frac{K\,\sigma_n\,(s-1)}{2}\right]^2 =$$

$$= K^2\,\sigma_n^2 \left[\frac{(s-1)\,(2s-1)}{6} - \frac{(s-1)^2}{4}\right] = \frac{K^2\,\sigma_n^2\,(s^2-1)}{12}\,. \tag{7.78}$$

Вычитание постоянной составляющей $K\sigma_n(s-1)/2$ из кодированного сигнала эквивалентно использованию s импульсов, амплитуды которых изменяются через $K\sigma_n$ в интервале от $-K\sigma_n(s-1)/2$ до $K\sigma_n(s-1)/2$. Таким образом, простым использованием как положительных, так и отрицательных импульсов мощность сигнала снижается на величину $[K\sigma_n(s-1)/2]^2$. В двоичном случае это означает использование двух импульсов с высотами $A/2$ и $-A/2$ вместо 0 и A. Средняя мощность биполярных импульсов равна, очевидно, $A^2/4$, а униполярных импульсов $A^2/2$.

Улучшение отношения сигнал/шум в случае сигнала с уменьшенной мощностью

$$\frac{S_o/N_o}{S_i'/N_i'} = \frac{48}{K^2\,(s^2-1)}\, s^{2B/f_m}. \tag{7.79}$$

Улучшение отношения сигнал/шум по напряжению определяется по формуле:

$$\frac{s_0/\sigma_0}{s_i'/\sigma_i'} = \frac{4\sqrt{3}}{K\sqrt{s^2-1}}\, s^{B/f_m} \,. \tag{7.80}$$

Заметим, что отношение сигнал/шум на входе приемника с учетом ф-лы (7.78)

$$\frac{S_i'}{N_i'} = \frac{S_i'}{\sigma_n^2} = \frac{K^2(s^2-1)}{12} \,. \tag{7.81}$$

7.6. Сравнение систем с кодированием и без кодирования

Выше были рассмотрены несколько широкополосных систем, которые позволяют обменивать отношение сигнал/шум на полосу частот. Примерами широкополосных систем без кодирования являются ЧМ (а также ФМ) и ФИМ. Кодово-импульсная модуляция является примером широкополосной системы с кодированием. При рассмотрении систем без кодирования (ЧМ и ФИМ) отмечалось, что улучшение отношения сигнал/шум по напряжению линейно зависит от полосы частот передачи [ф-лы (7.53) и (7.66б)]. В системе с кодированием (КИМ) улучшение отношения сигнал/шум по напряжению связано с полосой частот показательной зависимостью [ф-ла (7.80)]. Таким образом, обмен полосы частот на отношение сигнал/шум в системах с кодированием значительно более эффективен, чем в системах без кодирования.

В гл. 8 будет показано, что наибольшее теоретически возможное улучшение отношения сигнал/шум выражается как показательная зависимость от полосы частот. Таким образом, системы с кодированием весьма близки к реализации теоретически максимальной эффективности, которая вообще возможна. Нетрудно понять качественную сторону показательной зависимости в случае КИМ. В системах без кодирования при удвоении полосы частот передачи в два раза увеличивается отношение сигнал/шум по напряжению. В системе КИМ удвоение полосы частот позволяет в два раза увеличить число передаваемых импульсов. Если же число импульсов в комбинации возрастает в два раза, то число комбинаций возрастает в четыре раза. Например, из двух двоичных импульсов можно образовать четыре различных комбинации. Но из четырех импульсов можно образовать 16 комбинаций. Следовательно, при удвоении полосы частот число уровней квантования возрастает в четыре раза. Подобным же образом увеличение полосы частот в три раза позволяет во столько же раз увеличить число импульсов в комбинации. Но увеличение в три раза числа импульсов увеличивает в третьей степени число кодовых комбинаций. В общем случае расширение полосы частот в n

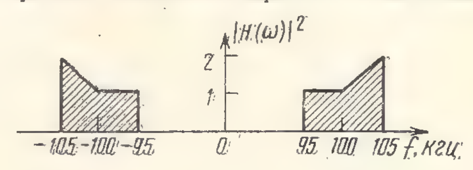
раз увеличивает число комбинаций по показательному закону (в n-й степени). Это позволяет по такому же закону увеличить число уровней квантования. Очевидно, шаг квантования Δu уменьшается по показательному закону с ростом полосы частот. Мощность шума квантования (или мощность выходного шума), равная $\Delta u^2/12$, уменьшается также по показательному закону. Это вызывает увеличение отношения сигнал/шум по показательному закону.

Самое важное преимущество системы с кодированием заключается в том, что она является наиболее подходящей для связей на большие расстояния с использованием множества ретрансляторов. При каждой ретрансляции сумма сигнала и шума «очищается» от шумов с помощью регенерации импульсов. Таким образом, каждый ретранслятор передает сигнал без шума. В действительности это означает, что в случае КИМ необходимо учитывать шум только в линии между соседними ретрансляторами. При передаче шумы не накапливаются. Таким образом, требуемая мощность сигнала при КИМ почти не зависит от общей длины линии связи. Действительно, КИМ позволяет вести передачу на любые расстояния с помощью достаточного числа ретрансляторов. Это основная причина для использования систем с кодированием. Наоборот, в системах без кодирования (или системах аналоговой модуляции) сигнал непрерывно подвергается воздействию шума и не может быть «очищен» или периодически регенерирован вдоль линии связи. Шумы на всем пути передачи накапливаются. Чем длиннее линия связи, тем больше мощность шума. Это ограничивает расстояние, на которое такие сигналы могут передаваться.

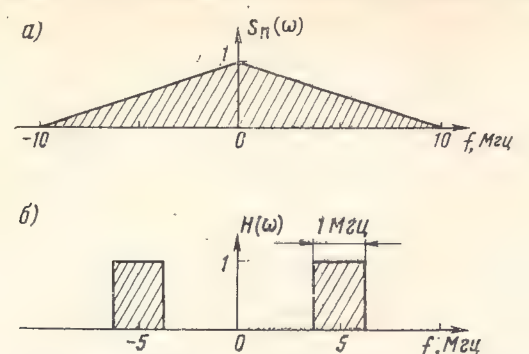
Читатель заметит аналогию между системами с кодированием и без кодирования с цифровыми и аналоговыми системами соответственно. В аналоговых системах (например, аналоговых вычислительных машинах) точность возрастает пропорционально напряжению, представляющему собой физическую переменную. Однако в цифровых системах при увеличении числа разрядов точность растет по показательному закону. Ошибки в аналоговых системах накапливаются, тогда как в цифровых системах (например, в цифровых вычислительных машинах) накопление не происходит и шум, лежащий ниже некоторого уровня, вообще не оказывает влияния на точность вычисления.

ЗАДАЧИ

7.1. Белый шум со спектром плотности мощности $N/2$, *вт/гц*, пропускается через полосовой фильтр, передаточная функция которого показана на рисунке. На выходе фильтра получается полосовой шум. Выразите это колебание через квадратурные составляющие по ф-ле (7.1). Определите спектры плотности мощности и средние значения квадратов составляющих $n_c(t)$ и $n_s(t)$.



7.2. Шумовое колебание, имеющее спектр плотности мощности $S_n(\omega)$, показанный на рисунке *(а)*, пропускается через идеальный полосовой фильтр



с передаточной функцией, показанной на рисунке *(б)*. Выразите выходное колебание фильтра через квадратурные составляющие по ф-ле (7.1).

7.3. Шум в некотором канале имеет равномерный спектр с плотностью мощности $S_n(\omega)=0{,}5\times10^{-3}$ *вт/гц*. По этому каналу передается сигнал ДБП-ПН. Модулирующее сообщение $f(t)$ имеет спектр, ограниченный частотой 5 *кгц*, несущая частота равна 100 *кгц*. Мощность модулированного сигнала (боковых полос) равна 10 *квт*. Входной сигнал в приемнике пропускается через идеальный полосовой фильтр, прежде чем поступает к детектору. Какой должна быть передаточная функция этого фильтра?

Чему равно отношение мощностей сигнала и шума на входе и выходе детектора?

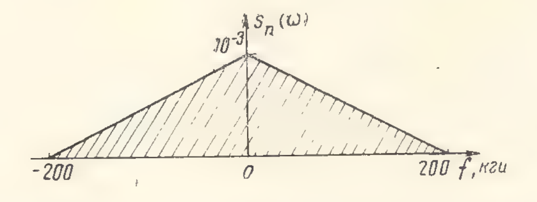
Найдите и изобразите графически спектр плотности мощности шума на выходе детектора.

7.4. По каналу, определенному в задаче 7.3, передается сигнал ОБП-ПН. Модулирующее сообщение имеет спектр, ограниченный частотой 5 *кгц*, несущая частота равна 100 *кгц*. Мощность модулированного сигнала равна 10 *квт*. До подачи на детектор входной сигнал пропускается через идеальный полосовой фильтр. Рассматривается передача верхней боковой полосы. Какой должна быть передаточная функция полосового фильтра на входе приемника?

Чему равно отношение мощностей сигнала и шума на входе и выходе детектора?

Как сравнить это отношение сигнал/шум с отношением сигнал/шум для ДБП-ПН из задачи 7.3? Дайте необходимые объяснения. Повторите эту задачу для случая передачи нижней боковой полосы. Дайте объяснения.

7.5. Спектр плотности мощности канального шума $S_n(\omega)$ показан на рисунке. По этому каналу передается сигнал ДБП-ПН. Спектр модулирующего



сообщения ограничен частотой 5 *кгц*, несущая частота равна 100 *кгц*. Мощность модулированного сигнала равна 10 *квт*. Сигнал на входе приемника пропускается через идеальный полосовой фильтр до подачи на детектор. Какую передаточную функцию должен иметь полосовой фильтр?

Чему равно отношение мощности сигнала к мощности шума на входе и выходе детектора?

Определите и изобразите графически спектр плотности мощности шума на выходе детектора.

7.6. Рассмотрите канал со спектром плотности мощности шума, показанным на рисунке к задаче 7.5. По этому каналу передается сигнал ОБП-ПН. Модулирующее сообщение, как и в задаче 7.5, имеет спектр, ограниченный частотой 5 *кгц*, а несущая частота равна 100 *кгц*. Мощность модулированного сигнала равна 10 *квт*. Сигнал на входе приемника пропускается через идеальный полосовой фильтр до подачи на детектор. Рассмотрите передачу верхней боковой полосы.

Какую передаточную функцию должен иметь полосовой фильтр на входе приемника?

Чему равно отношение мощности сигнала к мощности шума на входе и выходе детектора?

Как найденное отношение на выходе детектора сравнить с отношением сигнал/шум для системы ДБП-ПН из задачи 7.5?

Найдите и изобразите графически спектр плотности мощности шума на выходе детектора.

Повторите задачу для случая передачи нижней боковой полосы и приведите необходимые пояснения.

7.7. Канал с белым шумом (плотность мощности $S_n(\omega)=0{,}5\cdot 10^{-3}$ *вт/гц*) используется для передачи сигнала АМ. Спектр модулирующего сообщения ограничен частотой 5 *кгц*, несущая частота равна 100 *кгц*. Мощность боковых полос равна 10 *квт*, а амплитуда несущей — 200 *в* (мощность несущей 40 *квт*). Сигнал в приемнике пропускается через идеальный полосовой фильтр, а затем поступает на детектор огибающей. Определите отношение сигнал/шум на входе и выходе детектора.

7.8. Повторите задачу 7.7, если шум в канале имеет спектр плотности мощности, показанный на рисунке к задаче 7.5.

7.9. Сообщение $f(t)$ имеет среднее квадратическое значение $\overline{f^2(t)}=4/3$. Наибольшее положительное значение $f(t)$ равно 2 *в*, а наибольшее отрицательное —2 *в*. Известно также, что спектр $f(t)$ ограничен частотой 5 *кгц*. Это сообщение передается по каналу с белым шумом посредством ЧМ. Определите полосу частот, которую занимает передаваемый сигнал, если требуемое улучшение отношения сигнал/шум на выходе детектора равно 60.

Указание. Используйте ф-лу (7.48) и считайте, что $\Delta\omega=k_f|f(t)|_{\text{макс}}=2k_f$.

7.10. В сигнале широкополосной ЧМ амплитуда несущего колебания равна 100 *в*, а его частота—100 *Мгц*. Спектр модулирующей функции ограничен частотой 5 *кгц*, а ее средняя мощность равна 5000 *вт*. Постоянная модуляции $k_f=500\,\pi$, девиация частоты $\Delta f=75$ *кгц*. Шум в канале передачи имеет равномерный спектр с плотностью мощности $S_n(\omega)=10^{-3}$ *вт/гц*.

Определите передаточную функцию идеального полосового фильтра на входе приемника.

Найдите отношение сигнал/шум (по мощности) на входе и выходе детектора.

Определите и изобразите графически спектр шума на выходе детектора.

Найдите отношение сигнал/шум на выходе приемника, если сигнал передается посредством АМ, и сравните его с отношением сигнал/шум на выходе приемника ЧМ.

7.11. Рассмотрите систему ЧМ из задачи 7.10. Решено использовать предыскажения с последующей коррекцией для снижения шума на выходе приемника. В предыскажающем фильтре частоты ω_1 и ω_2 равны соответственно 1,5 *кгц* и 30 *кгц*. Найдите отношение сигнал/шум на выходе детектора. Сравните его с отношением сигнал/шум в отсутствие предыскажений из задачи 7.10.

7.12. Улучшение отношения сигнал/шум при фазово-импульсной модуляции определяется по ф-ле (7.66а). Это улучшение пропорционально $k_1^2\overline{f^2(t)}B$. Та-

ким образом, его можно сделать большим, если увеличивать B, $\overline{f^2(t)}$ или k_1. Если увеличивается B, то расширяется полоса частот передаваемого сигнала, при увеличении $\overline{f^2(t)}$ растет средняя мощность, т. е. отношение сигнал/шум улучшается за счет соответствующего изменения характеристик сигнала. Какие характеристики сигнала изменяются, если улучшение достигается за счет увеличения k_1? Можно ли бесконечно увеличивать k_1 и улучшать отношение сигнал/шум? Из ф-лы (7.66а) следует, что отношение сигнал/шум можно также увеличить, если уменьшить α. Какая характеристика сигнала изменяется при уменьшении α? Можно ли α уменьшать бесконечно? Обсудите эти вопросы. Рассмотрите также, что произойдет, если полоса пропускания канала сильно уменьшится.

7.13. Модулирующее сообщение $f(t)$ передается посредством ФИМ. Задано $|f(t)|_{\text{макс}}=2$ *в*, $\overline{f^2(t)}=4/3$. Спектр $f(t)$ ограничен частотой 5 *кгц*. Производится выборка с частотой 10 000 *имп/сек* и полученные отсчеты передаются посредством ФИМ. Полоса пропускания канала 300 *кгц*, длительность импульсов $\tau=20$ *мксек*. Чему равно наибольшее допустимое значение постоянной k_1? Для этого значения k_1 определите улучшение отношения сигнал/шум.

7.14. Улучшение отношения сигнал/шум (по мощности) в ф-ле (7.72) получено для однополярных двоичных импульсов, т. е. импульсов, которые принимают значения 0 и $K\sigma_n$. Получите формулу для улучшения отношения сигнал/шум, если используются биполярные импульсы с амплитудами $-K\sigma_n/2$ и $K\sigma_n/2$.

7.15. Рассмотрите вновь сообщение $f(t)$ из задачи 7.13 с $|f(t)|_{\text{макс}}=2$ *в* и $\overline{f^2(t)}=4/3$. Спектр $f(t)$ ограничен частотой 5 *кгц* и выборка производится с частотой 10 000 *имп/сек*. Эти выборочные значения квантуются и кодируются двоичным кодом. Шаг квантования равен 1/32 *в*. Эффективное значение шума в канале 0,1 *в*. Предполагается, что значения $f(t)$ имеют равномерное распределение в интервале от -2 до $+2$ *в*. Для передачи используются однополярные импульсы с амплитудами 0 и $K\sigma_n$, где $K=10$. Определите ширину спектра сигнала, отношение сигнал/шум на входе и выходе приемника и улучшение отношения сигнал/шум.

Какая потребовалась бы полоса частот, чтобы такое же улучшение получить при ЧМ?

Повторите задачу для биполярных импульсов с амплитудами $-K\sigma_n/2$ и $K\sigma_n/2$ $(K=10)$.

Сравните эту систему передачи с ФИМ из задачи 7.13.

7.16. Повторите задачу 7.15, если вместо двоичных импульсов используются четверичные, а шаг квантования установлен 1/64 *в*. Проведите вычисления для случаев однополярных и биполярных импульсов. $K=10$.

ГЛАВА

8

Введение в теорию передачи информации

8.1. Мера информации

Каждое сообщение переносит некоторую информацию, однако одни сообщения переносят больше информации, чем другие. Детальный анализ показывает, что вероятность появления события тесно связана с количеством информации. Так, в сообщении о высоковероятном событии содержится мало информации в сравнении с той, которая содержалась бы в событии маловероятном. Фактор неожиданности или неопределенности события, очевидно, пропорционален количеству информации. Если кто-то говорит, что солнце восходит на востоке, он не доставляет абсолютно никакой информации, поскольку всякий знает, что это верно. В таком событии, как ежедневный восход солнца на востоке, нет никакой неопределенности. Иными словами, вероятность этого события равна единице. Однако если служба погоды сообщит, что в январе температура воздуха достигнет 70°C, то в этом сообщении будет содержаться огромное количество информации. Это объясняется тем, что такое событие является полностью неожиданным и вероятность его появления ничтожно мала ($P \to 0$). Другими словами, это событие чрезвычайно неопределенно, и поэтому количество содержащейся в нем информации стремится к очень большой величине. Неожиданность события, разумеется, является результатом неопределенности или непредвиденности. Чем больше непредвиденность события, тем больше неожиданность и, следовательно, тем больше информация. Вероятность события является мерой его неожиданности и потому связана с информационной содержательностью события.

Итак, количество информации, полученной в результате появления некоторого события, является величиной, обратной вероятности появления этого события. Какой должна быть природа этой взаимосвязи? Очевидно, если событие достоверно (вероятность равна единице), оно доставляет нулевое количество информации. С другой стороны, если событие невозможно (вероятность равна нулю), то его появление доставляет бесконечное количество информации. Это наводит на мысль, что количество информации должно быть логарифмической функцией величины, обратной вероятности события;

$$I \sim \log \frac{1}{P};\qquad (8.1)$$

где P — вероятность появления события, а I — количество информации, полученной с появлением данного события.

Покажем теперь, что с инженерной точки зрения содержащаяся в событии информация совпадает с той, которая получена при качественном подходе (8.1). Что понимается под инженерной точкой зрения? Инженер интересуется тем, как наиболее эффективно передать информацию, содержащуюся в сообщениях. С его точки зрения количество информации в сообщении пропорционально времени, затрачиваемому на передачу сообщения. Ниже будет показано, что такая концепция информации также приводит к ф-ле (8.1), которая означает, что на передачу сообщения, имеющего более высокую определенность (или бóльшую вероятность), требуется меньшее время, чем на передачу менее вероятного сообщения. В этом легко убедиться на примере передачи английского текста посредством кода Морзе.

Код Морзе образуется различными комбинациями двух символов (таких, как «посылка» и «пауза» или импульсов высотой a и $-a$, *в*). Каждая буква представляется определенной комбинацией этих символов, имеющей определенную длину. Очевидно, короткие кодовые комбинации приписываются буквам *e, t, a* и *o*, которые встречаются более часто. Длинные кодовые комбинации приписываются буквам *x, k, q* и *z*, которые встречаются более редко. Каждая из букв может рассматриваться как сообщение. Очевидно, что буквы с высокой определенностью (с высокой вероятностью) нуждаются в меньшем времени передачи (так как кодовые комбинации короче), чем те, которые имеют меньшую вероятность. Покажем теперь, что минимально необходимое время для передачи символа (или сообщения) с вероятностью P действительно пропорционально $\log(1/P)$. Предположим, что требуется передать любое из двух равновероятных сообщений — a или b. Ими могут быть, например, сообщения о погоде, такие, как «солнечно» и «дождливо». Каждое из этих сообщений можно передавать подходящим колебанием. Предполагая, что для передачи упомянутых сообщений применяются двоичные импульсы, можно условиться, что отсутствие импульса (импульс высотой 0 *в*) соответствует сообщению a («солнечно»), а наличие импульса (импульс высотой 1 *в*) — сообщению b («дождливо»).

Очевидно, как минимум потребуется один двоичный импульс, чтобы передать любое из двух равновероятных сообщений. Информация в любом из них соответственно определяется как 1 *дв. ед.* (двоичная единица). Заметим, что длина сообщения не влияет на его информационную содержательность. Для передачи одного из двух возможных равновероятных сообщений вне зависимости от их длины или какой-нибудь другой характеристики

требуется один двоичный импульс. Очевидно, один двоичный импульс может передать одну двоичную единицу информации.

Рассмотрим далее случай четырех равновероятных сообщений. Если эти сообщения передаются двоичными импульсами, то на каждое сообщение потребуется группа из двух двоичных импульсов. Каждый двоичный импульс может принимать два состояния, поэтому совокупность двух импульсов образует четыре различные комбинации, которые и приписываются каждому из четырех сообщений (табл. 8.1). Следовательно, для передачи любого из

Таблица 8.1

ПЕРЕДАЧА ЧЕТЫРЕХ РАВНОВЕРОЯТНЫХ СООБЩЕНИЙ ПРИ ПОМОЩИ ДВОИЧНЫХ ИМПУЛЬСОВ

Символ сообщения	Двоичное число	Сигнал двоичного кода	Четверичное число	Сигнал четверичного кода
A	00		0	0 в
B	01		1	1 в
C	10		2	2 в
D	11		3	3 в

четырех равновероятных сообщений требуется два двоичных импульса. Передача каждого из этих сообщений занимает в два раза большее время, чем то, которое требуется для передачи одного из двух равновероятных сообщений, а значит, сообщение содержит в два раза больше информации, т. е. 2 *дв. ед.*

Подобным образом комбинацией из трех двоичных импульсов можно передать любое из восьми равновероятных сообщений, так как эти три импульса образуют восемь различных комбинаций. Следовательно, каждое из восьми равновероятных сообщений содержит 3 *дв. ед.* информации. В общем случае любое из n равновероятных сообщений содержит $\log_2 n$ *дв. ед.* информации в соответствии с инженерным подходом к определению информации. Это означает, что для передачи любого такого сообщения потребуется как минимум $\log_2 n$ двоичных импульсов. Заметим, что вероятность появления любого из этих событий $P=1/n$. Следовательно,

$$I=\log_2 n=\log_2 (1/P). \tag{8.2}$$

Этот результат получен для весьма специфического случая равновероятных сообщений. Можно показать также, что даже если сообщения не равновероятны, то в среднем на передачу со-

общения, имеющего вероятность появления P, требуется $\log_2(1/P)$ двоичных импульсов.

Для доказательства этого рассмотрим источник, который создает сообщения $a_1, a_2, \ldots, a_n$ с вероятностями $P_1, P_2, \ldots, P_n$. Пусть источник генерирует последовательность из N сообщений. Если N становится очень большим, то по закону больших чисел эта последовательность будет содержать сообщение a_1 NP_1 раз, сообщение a_2 NP_2 раз и т. д. Поскольку появление каждого сообщения независимо, вероятность $P(S)$ появления любой последовательности S из N сообщений $P(S) = (P_1)^{NP_1}(P_2)^{NP_2}\ldots(P_n)^{NP_n}$.

Заметим, что, поскольку N очень велико, каждая из возможных последовательностей содержит одинаковое число сообщений $a_1(NP_1)$, $a_2(NP_2)$ и т. д. Следовательно, все эти последовательности имеют одинаковую вероятность появления $P(S)$, определяемую приведенной выше формулой. Одна такая последовательность содержит информацию

$$I(S) = \log_2 \frac{1}{P(S)} = N \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i}.$$

Поскольку последовательность S составлена из N сообщений, то средняя информация на сообщение

$$\frac{I(S)}{N} = \sum_{i=1}^{n} P_i \log_2 \frac{1}{P_i}, \quad \text{дв. ед.} \tag{8.2a}$$

Это и есть средняя информация на сообщение. Поскольку сообщения $a_1, a_2, \ldots, a_n$ появляются с вероятностями $P_1, P_2, \ldots, P_n$, то, очевидно, сообщение a_k, появляющееся с вероятностью P_k, переносит информацию, равную $\log_2 \frac{1}{P_k}$, что и является желаемым результатом.

Из изложенного очевидно, что информационная мера сообщений (в двоичных единицах) равна минимальному числу двоичных импульсов, необходимых для кодирования этого сообщения. Такое определение информации может показаться на первый взгляд ограниченным, поскольку оно применимо только к информации дискретной природы, такой, как передача некоторого дискретного и конечного числа символов или сообщений. Однако важным результатом теории информации является то, что любая подлежащая передаче информация может быть всегда представлена в двоичной форме без потери общности. В предыдущей главе было показано, что содержащуюся в непрерывном сигнале с ограниченным спектром информацию можно представить дискретным числом отсчетов в секунду. Затем можно представить эти отсчеты кодом, в котором используются двоичные импульсы.

Ниже будет показано, что каждая система связи (или канал) способна передать определенное количество информации в секунду. Это предельное количество информации называется емкостью C канала. Таким образом, данный канал может передать в одну секунду количество информации, не превосходящее C, *дв. ед./сек.* Емкость канала зависит от полосы пропускания и отношения мощностей сигнала и шума в канале.

Вместо двоичных импульсов при кодировании можно использовать M-ичные импульсы (импульсы, которые могут принимать M различных значений). Покажем, что каждый M-ичный импульс может переносить $\log_2 M$, *дв. ед.* информации. Для этого допустим, что требуется передать любое не из двух, а из четырех равновероятных сообщений. Очевидно, передать такую информацию одним двоичным импульсом невозможно, так как он может принимать лишь два состояния. Если же используются четверичные импульсы, высота которых может принимать четыре значения, например 0, 1, 2 и 3 *в*, то любое из четырех сообщений можно передать одним импульсом (табл. 8.1). Следовательно, один четверичный импульс может передать информацию, содержащуюся в двух двоичных импульсах, т. е. он содержит 2 *дв. ед.* информации. Подобным образом для передачи любого из восьми возможных сообщений требуется группа из трех двоичных импульсов на каждый символ. С другой стороны, такое сообщение может быть передано одним восьмеричным импульсом, который может принимать восемь состояний или значений. Следовательно, один восьмеричный импульс переносит 3 *дв. ед.* информации. Легко видеть, что импульс, который может принимать M различных состояний или M различных уровней, переносит $\log_2 M$, *дв. ед.* информации.

Итак, чем большее число различных уровней может принимать импульс, тем больше информация, переносимая каждым импульсом. Импульс, который может принимать бесконечное число различных уровней, переносит бесконечное количество информации. Это означает, что возможно передать любое количество информации единственным импульсом, который может принимать бесконечное число различных состояний. Этот результат является несколько неожиданным, однако он вполне справедлив и логичен. Если импульс может принимать бесконечное число различных уровней, то, как и раньше, любому сообщению независимо от его длины соответствует один из возможных уровней импульса. Например, можно считать, что какой-либо уровень импульса представляет полное содержание данной книги. Если требуется передать содержание книги, то необходимо передать всего один импульс соответствующего уровня. Перечисление всей таблицы соответствий в этом случае невозможно, тем не менее этот пример показывает возможность передачи бесконечного количества информации одним импульсом.

В связи с изложенным возникает вопрос: почему не используются импульсы, которые принимают бесконечное множество различных уровней? Существуют ограничения, накладываемые на систему практическими причинами. Следует помнить, что во всех рассуждениях о передаче информации посредством импульсов имеется в виду вся система (включая передатчик и приемник). Следовательно, чтобы передать определенную информацию, необходимо не только передать, но и принять импульсы. Более того, необходимо обеспечить распознавание различных уровней им-

пульсов. Что же мешает передавать импульсы с бесконечным числом различных состояний? В реальных условиях импульсы должны иметь конечные амплитуды, а это значит, что при бесконечном числе состояний интервал между ними становится бесконечно малым. Так как в любом реальном канале присутствует шум, невозможно различить уровни, отличающиеся менее чем на амплитуду шума. Итак, из-за наличия шума в канале уровни сигнала должны отличаться, по крайней мере, на величину амплитуды шума.

8.2. Емкость канала

Как уже указывалось, полоса частот и мощность шума накладывают ограничения на скорость передачи информации по каналу. Можно строго показать, что по каналу, в котором действует гауссов шум, можно передавать информацию с наибольшей скоростью *C*, *дв. ед.* в секунду, где *C* — емкость канала, определяемая соотношением

$$C = B \log_2\left(1 + \frac{S}{N_к}\right), \tag{8.3}$$

B — ширина полосы частот в герцах, S — мощность сигнала, $N_к$ — мощность шума.

Выражение (8.3) справедливо только для белого гауссова шума. Для других видов шума емкость канала выражается другими формулами. Строгое доказательство ф-лы (8.3) выходит за рамки данной книги[1]). Здесь приводится упрощенное доказательство, основанное на предположении, что если сигнал смешан с шумом, то амплитуда сигнала может быть измерена лишь с точностью до эффективного значения напряжения шума. Другими словами, неопределенность оценки точного значения амплитуды сигнала равна квадратному корню из среднего квадрата шумового напряжения.

Обозначим средние мощности сигнала и шума через S и $N_к$ соответственно. Если предположить, что сопротивление нагрузки равно 1 *ом*, то корень из среднего квадрата принимаемого сигнала равен $\sqrt{S+N_к}$, *в*, а корень из среднего квадрата напряжения шума равен $\sqrt{N_к}$, *в*. Попробуем теперь различить принимаемый сигнал с амплитудой $\sqrt{S+N_к}$, *в*, в присутствии шума с амплитудой $\sqrt{N_к}$, *в*. Как следует из указанных выше предположений, изменения входного сигнала, меньшие чем $\sqrt{N_к}$, приемник не различает. Следовательно, число уровней, которое может быть различимо без ошибок, определится из выражения

$$M = \frac{\sqrt{S+N_к}}{\sqrt{N_к}} = \sqrt{1 + \frac{S}{N_к}}. \tag{8.4}$$

---

[1]) См. К. Шеннон. Работы по теории информации и кибернетике. Издательство иностр. лит., 1963.—*Прим. ред.*

Итак, максимальное число M определяется выражением (8.4). Наибольшее количество информации, переносимое каждым импульсом, имеющим $\sqrt{1+S/N_к}$ различных уровней,

$$I = \log_2 \sqrt{1+\frac{S}{N_к}} = \frac{1}{2}\log_2\left(1+\frac{S}{N_к}\right), \text{ дв. ед.} \tag{8.5}$$

Теперь можно определить емкость канала. Емкость канала — это максимальное количество информации, которое может быть передано по каналу в секунду. Если канал способен передать максимум K импульсов в секунду, очевидно, его емкость

$$C = \frac{K}{2}\log_2\left(1+\frac{S}{N_к}\right), \text{ дв. ед./сек.} \tag{8.6}$$

В гл. 5 было показано, что в полосе частот системы nf_m, *гц*, можно передавать $2nf_m$ независимых импульсов в секунду. При этих условиях принятый сигнал дает точное значение амплитуды импульсов, но не воспроизводит в деталях их форму. Так как интерес в данном случае представляет только амплитуда импульсов, а не их форма, то это означает, что по системе с полосой частот B, *гц*, можно передавать импульсы с максимальной скоростью $2B$ импульсов в секунду. Поскольку каждый импульс может переносить максимум $\frac{1}{2}\log_2(1+S/N_к)$, *дв. ед.*, информации, по системе с полосой частот B можно передавать информацию с максимальной скоростью

$$C = B\log_2\left(1+\frac{S}{N_к}\right), \text{ дв. ед./сек.} \tag{8.7}$$

Емкость канала ограничивается, таким образом, шириной полосы частот канала (или системы) и шумом. Для канала без шума $N_к = 0$ и емкость канала становится бесконечной. На практике, однако, $N_к$ всегда конечно и, следовательно, емкость канала конечна [1].

---

[1]) Это верно, если даже ширина полосы B бесконечна. Шум является белым, с равномерной спектральной плотностью в пределах всего частотного диапазона. Следовательно, если полоса частот B увеличивается, то увеличивается и $N_к$; емкость канала остается конечной, если даже $B = \infty$.

Если $N/2$ — спектральная плотность мощности, то $N_к = NB$ и

$$C = B\log_2\left(1+\frac{S}{NB}\right),$$

$$\lim_{B\to\infty} C = \frac{S}{N}\frac{NB}{S}\log_2\left(1+\frac{S}{NB}\right).$$

Последний предел можно найти, вспомнив, что

$\lim\limits_{x\to 0}\frac{1}{x}\log_2(1+x) = \log_2 e = 1{,}44$. Следовательно,

$$\lim_{B\to\infty} = \frac{S}{N}\log_2 e = 1{,}44\frac{S}{N}.$$

Соотношение (8.7) известно как формула Хартли-Шеннона и считается основной теоремой теории информации[1]). Из этой теоремы следует, что полоса частот и мощность сигнала могут обмениваться друг на друга. Чтобы передавать информацию с заданной скоростью, можно снизить мощность передаваемого сигнала, если соответственно увеличить полосу частот. Подобным образом можно сократить полосу частот за счет увеличения мощности сигнала. Как было установлено ранее, модуляция позволяет осуществить этот обмен между полосой частот и отношением сигнал/шум. В свете этой теоремы легко объясняется улучшение отношения сигнал/шум в широкополосных системах ЧМ и КИМ.

Следует, однако, помнить, что емкость канала представляет собой максимальное количество информации, которое может быть передано по каналу в секунду. Чтобы достигнуть такой скорости передачи, информация должна быть соответствующим образом обработана или закодирована наиболее эффективно. Утверждение, что такое кодирование возможно, является важнейшим результатом созданной Шенноном теории информации. К настоящему времени, однако, используемые системы связи (системы без кодирования такие, как АМ, ЧМ и т. д.) не достигают этой максимальной скорости.

8.3. Передача непрерывных сигналов

Проиллюстрируем смысл формулы Хартли-Шеннона применительно к обмену полосы частот и отношения сигнал/шум для непрерывных сигналов, спектр которых ограничен частотой f_m, *гц*. Из теоремы отсчетов (§ 1.15) следует, что информация в таком сигнале полностью определяется $2f_m$ отсчетами в секунду. Следовательно, чтобы передать содержащуюся в сигнале информацию, необходимо передать только эти дискретные отсчеты.

Очередной важный вопрос состоит в следующем: какое количество информации содержится в каждом отсчете? Количество информации зависит от того, сколько различных уровней или значений могут принимать отсчеты. В действительности отсчеты могут принимать любое значение и, следовательно, для их передачи потребуются импульсы, способные принимать бесконечное число различных уровней. Очевидно, информация, переносимая каждым отсчетом, составляет бесконечное число двоичных единиц. Таким образом, информация, содержащаяся в непрерывном сигнале с ограниченным спектром, равна бесконечности.

При наличии шума (с конечным значением $N_к$) емкость канала конечна. Поэтому передать всю информацию, содержащуюся в

---

[1]) Теория информации есть совокупность результатов, базирующихся на количественном определении информации; она составляет часть более широкой области — статистической теории связи, включающей все вопросы вероятностного анализа систем связи. См., например, Р. Фано. Передача информации. Статистическая теория связи. Изд-во «Мир», 1965.

сигнале с ограниченным спектром по реальному каналу с шумом невозможно. В отсутствие шума ($N_к = 0$) емкость канала бесконечна и по нему может быть передан любой полезный сигнал. Передать всю информацию, содержащуюся в непрерывном сигнале, при наличии шума невозможно, если мощность передаваемого сигнала остается конечной. Из-за шума в принимаемом сигнале всегда имеется какая-то неопределенность. Передача полной информации означала бы нулевую величину неопределенности. В самом деле, величина неопределенности может быть сделана сколь угодно малой путем увеличения емкости канала (расширением полосы частот и (или) увеличением мощности сигнала), однако она никогда не может стать равной нулю.

Важно заметить, что неопределенность вносится шумом канала, а не передатчиком. Следовательно, можно передать всю информацию, содержащуюся в непрерывном сигнале, но восстановить ее в приемнике практически невозможно. Количество информации, которое может быть воспроизведено в секунду в приемнике, не превышает *C, дв. ед./сек.* Именно так происходит при непосредственной передаче непрерывных сигналов (например, посредством АМ и ЧМ). В этих случаях передатчиком передается полная информация, но, поскольку канал имеет конечную емкость, только *C, дв.ед./сек* информации воспроизводится приемником.

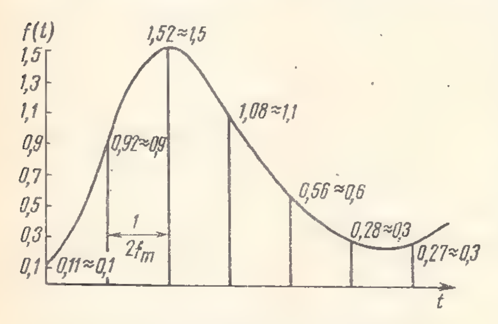


Рис. 8.1. К пояснению квантования непрерывного сообщения

Существует и другой способ передачи информации. Он заключается в том, что передаваемый сигнал аппроксимируют, уменьшая его информационную содержательность до значения *C, дв.ед./сек*, а затем передают этот аппроксимированный сигнал, который имеет конечную информационную содержательность. Теперь приемник в состоянии воспроизвести всю информацию, которая была передана. В этом как раз и состоит сущность кодово-импульсной модуляции.

Приближенное представление непрерывного сигнала, как сигнала с конечной информационной содержательностью в секунду, достигается посредством квантования, описанного в гл. 7. Рассмотрим непрерывный сигнал с ограниченной полосой частот f_m, *гц* (рис. 8.1). Чтобы передать содержащуюся в этом сигнале информацию, необходимо передать только $2f_m$ отсчетов в секунду. Отсчеты также показаны на этом рисунке. Как отмечалось ранее, отсчеты могут принимать любые значения. Для их непосредственной передачи требуются импульсы, которые могут принимать бесконечное число уровней. Так как из-за шума невозможно точно

восстановить высоту этих импульсов в приемнике, их амплитуды округляются до ближайших значений из конечного числа разрешенных уровней. На рис. 8.1 амплитуды импульсов округляются до ближайших значений, следующих через одну десятую вольта. Из рисунка видно, что каждый передаваемый импульс представляется одним из 16 уровней и, следовательно, переносимая им информация составляет $\log_2 16=4$ *дв. ед.* Поскольку имеется $2f_m$ отсчетов в секунду, общая информационная содержательность квантованного сигнала равна $8f_m$, *дв. ед./сек.* Если емкость канала больше или равна $8f_m$, *дв. ед./сек,* то вся передаваемая информация будет полностью восстановлена без какой-либо неопределенности. Это означает, что принятый сигнал будет точной копией переданного квантованного сигнала.

Возникает вопрос, может ли шум, вносимый в процессе передачи, вызвать дополнительную неопределенность и тем самым увеличить общую неопределенность амплитуды принимаемого сигнала сверх 0,1 *в?* Можно показать, что если емкость канала равна $8f_m$, *дв. ед./сек,* то в процессе передачи дополнительная неопределенность не вносится[1]. Предположим, что для передачи импульсов используется канал с полосой частот f_m, *гц.* Поскольку требуемая емкость канала равна $8f_m$, *дв. ед./сек,* необходимое отношение сигнал/шум по мощности определяется из соотношения

$$8f_m=f_m\log_2[(S+N_к)/N_к]. \tag{8.8}$$

Следовательно, $(S+N_к)/N_к=256$.

Число различимых приемником уровней равно $\sqrt{(S+N_к)/N_к}$. Следовательно, в данном случае приемник может различить без ошибки 16 уровней. Хотя в процессе передачи на сигнал накладывается некоторый шум, уровни отстоят друг от друга настолько, чтобы быть различимыми приемником, т. е. канал емкостью $8f_m$, *дв. ед./сек.* может обеспечить передачу информации $8f_m$, *дв. ед./сек* практически без ошибок.

8.4. Обмен полосы частот на отношение сигнал/шум

Данный сигнал может быть передан с заданной величиной неопределенности по каналу с ограниченной емкостью. Требуемую емкость канала можно получить различными комбинациями полосы частот и мощности сигнала. Действительно, один из показателей можно обменять на другой. Покажем, как осуществляется такой обмен.

Рассмотрим передачу сигнала $f(t)$ (рис. 8.1). Если допустимая неопределенность равна 0,1 *в,* то информационная содержа-

---

[1]) Это верно в предположении, что приемник безошибочно различает уровни сигналов, отличающиеся на $\sqrt{N_k}=\sigma_n$. В действительности такие сигналы принимают с большой вероятностью ошибки, и чтобы снизить ее до незначительной величины, требуется интервал между соседними уровнями сигнала $k\sigma_n$, где $k\gg 1$. (Например, чтобы вероятность ошибки была порядка 10^{-6}, требуется $k=10$) — *Прим. ред.*

тельность сигнала составляет 8 f_m, *дв. ед./сек.* Теперь покажем, что эта информация может быть передана при различных комбинациях полосы частот и мощности сигнала.

Один из возможных способов передачи состоит в том, чтобы посылать непосредственно $2f_m$ отсчетов в секунду. Каждый отсчет может принимать любое из 16 состояний (шестнадцатиричный импульс). В этом случае отношение сигнал/шум должно быть таким, чтобы можно было различить 16 состояний. Очевидно, $\sqrt{(S+N_к)/N_к}=16$. Далее, для передачи $2f_m$ импульсов в секунду требуется полоса частот f_m, *гц*. Следовательно, необходимая емкость канала C определится в соответствии с ф-лой (8.7):

$$C=f_m\log_2[(S+N_к)/N_к]=f_m\log_2 16^2=8f_m,\ дв.\ ед./сек.$$

Итак, емкость канала в точности равна количеству информации в секунду, содержащейся в сигнале $f(t)$. Другой способ состоит в передаче отсчетов рис. 8.1 с помощью четверичных импульсов (импульсов, которые могут принимать четыре состояния). Очевидно, что для передачи каждого отсчета, который может принимать 16 состояний, понадобится группа из двух четверичных импульсов. Теперь отношение сигнал/шум, требуемое для различения в приемнике импульсов, принимающих четыре различных состояния, равно $\sqrt{(S+N_к)/N_к}=4$. При таком способе передачи необходимая мощность сигнала уменьшается. Однако теперь нужно передавать вдвое большее число импульсов в секунду, т. е. $4f_m$. Следовательно, потребуется полоса частот $2f_m$, *гц*. Емкость канала C в данном случае:

$$C=2f_m\log_2[(S+N_к)/N_к]=2f_m\log_2 4^2=8f_m,\ дв.\ ед./сек.$$

Из этого примера видно, что заданное количество информации может быть передано при различных комбинациях мощности сигнала и полосы частот, причем один показатель может обмениваться на другой. Сигнал $f(t)$ можно передать также двоичными импульсами (8 f_m импульсов в секунду), которые потребуют $\sqrt{(S+N_к)/N_к}=2$ и ширину полосы частот $4f_m$, *гц*. Интересно заметить, что сигнал $f(t)$ можно передать по каналу с меньшей чем f_m полосой частот, если имеется достаточная мощность передатчика (см. задачу 8.3).

Следует обратить внимание на то, что обмен полосы частот и мощности сигнала не происходит автоматически. Для этого нужно видоизменить или преобразовать содержащуюся в сигнале информацию так, чтобы занять требуемую полосу частот. На практике такое преобразование осуществляется посредством различных видов модуляции. Однако следует подчеркнуть, что не каждая система связи полностью использует емкость, определяемую ее полосой частот и мощностью сигнала. Одни виды модуляции оказываются лучше по использованию емкости канала, другие — хуже. В гл. 7 было показано, что системы с кодированием предпоч-

тительней систем без кодирования (таких как ЧМ, ФИМ) с точки зрения обмена полосы частот на отношение $S/N_к$.

Из формулы Хартли—Шеннона можно вывести идеальный закон обмена полосы частот на отношение $S/N_к$. Рассмотрим сообщение, которое имеет полосу частот f_m, *гц*. Предположим, что содержащаяся в этом сигнале информация равна I, *дв. ед./сек.* Пусть, далее, сигнал так кодируется (или модулируется), что его полоса частот становится равной B, *гц*. Модуляция сигнала не изменяет его информационной содержательности. Модулированный сигнал поступает затем на вход детектора (рис. 8.2). Пусть мощность сигнала и мощность шума на входе детектора равны соответственно S_i и N_i. Очевидно,

$$I = B \log_2\left(1 + \frac{S_i}{N_i}\right). \qquad (8.9)$$

Рис. 8.2. К изменению отношения сигнал/шум при детектировании сигнала

На выходе детектора получается исходный сигнал $f(t)$ с полосой частот f_m и некоторый шум. Обозначим мощность сигнала и мощность шума на выходе через S_o и N_o соответственно. В случае идеального детектора информация I выходного и входного сигналов одинакова. Следовательно,

$$I = f_m \log_2\left(1 + \frac{S_o}{N_o}\right). \qquad (8.10)$$

Из ф-л (8.9) и (8.10) получаем

$$B \log_2\left(1 + \frac{S_i}{N_i}\right) = f_m \log_2\left(1 + \frac{S_o}{N_o}\right),$$

$$1 + \frac{S_o}{N_o} = \left(1 + \frac{S_i}{N_i}\right)^{B/f_m}. \qquad (8.11)$$

На практике S_o/N_o и $S_i/N_i \gg 1$, поэтому

$$S_o/N_o \approx (S_i/N_i)^{B/f_m}. \qquad (8.12)$$

Если рассматривается отношение сигнал/шум по напряжению, а не по мощности, то

$$s_o/\sigma_o \approx \left(\frac{s_i}{\sigma_i}\right)^{B/(2f_m)}. \qquad (8.13)$$

Итак, в идеальной системе отношение мощности сигнала к мощности шума на выходе (S_o/N_o) растет по показательному закону с увеличением полосы частот B. Тот же результат сохраняется для отношения сигнал/шум по напряжению. Эта характеристика, очевидно, намного превосходит то, что дают некодированные широкополосные системы, такие, как ЧМ и ФИМ, рассмотренные в гл. 7. Для этих систем, как было показано, отношение мощности сигнала к мощности шума растет пропорционально квадрату по-

лосы частот B. Иное наблюдается в системах с кодированием (КИМ), в которых отношение сигнал/шум действительно улучшается по показательному закону с расширением полосы частот B [ур-ние (7.79)].

Пример 8.1.

В этом примере, пользуясь понятиями теории информации, рассчитаем полосу частот видеосигнала (сигнала изображения) в телевидении.

Телевизионное изображение можно рассматривать состоящим из 300 000 мелких элементов изображения. Каждый из этих элементов может принимать 10 различимых градаций яркости. Допустим, что все градации яркости равновероятны и за одну секунду передается 30 кадров изображения. Кроме того, пусть известно, что для удовлетворительного воспроизведения изображения необходимое отношение сигнал/шум равно 1000 (30 *дб*).

Располагая этими данными, рассчитаем ширину полосы частот, необходимую для передачи видеосигнала. Поскольку каждый элемент изображения может принимать 10 уровней с равной вероятностью, то один элемент изображения содержит $\log_2 10 = 3{,}32$ *дв. ед./элемент*, один кадр изображения содержит $300\,000 \times 3{,}32 = 996\,000$ *дв. ед./кадр*. Так как в одну секунду передается 30 кадров, то количество информации в секунду равно $996\,000 \times 30 = 29{,}9 \times \times 10^6$ *дв.ед./сек*. Для передачи этой информации необходима емкость канала $C = 29{,}9 \cdot 10^6$ *дв. ед./сек*. По ф-ле (8.7) при $S/N_k = 1000$ находим

$$C = 29{,}9 \cdot 10^6 \approx B \log_2 1000 = 9{,}95\, B.$$

Поэтому $B = 3{,}02 \cdot 10^6$ *гц* ≈ 3 *Мгц*.

8.5. Эффективность систем кодово-импульсной модуляции

Выше отмечалось, что отношение сигнал/шум на выходе приемника при КИМ с расширением полосы частот растет по показательному закону, как и в идеальной системе. Однако характеристики КИМ еще далеки от тех, которые предсказывает формула Хартли—Шеннона. Рассмотрим s-ичную систему КИМ. Сообщение $f(t)$ имеет полосу частот f_m, *гц*, поэтому отсчеты следуют с частотой $2f_m$ в секунду. Пусть число уровней квантования равно M. Предполагая все уровни равновероятными, для количества информации на отсчет получим $\log_2 M$, *дв. ед*. Таким образом, полная информация, содержащаяся в сигнале, равна $2 f_m \log_2 M$, *дв. ед./сек*. Чтобы передать такой сигнал, теоретически необходим канал, емкость которого

$$C = 2 f_m \log_2 M, \text{ дв. ед./сек.} \tag{8.14}$$

Если для передачи используются s-ичные импульсы (импульсы, которые могут принимать s состояний), то для представления одного отсчета потребуется группа из $\log_s M$ импульсов. Следовательно, всего необходимо передать $2 f_m \log_s M$ s-ичных импульсов в секунду. Для этого требуется полоса частот

$$B = f_m \log_s M. \tag{8.15}$$

Из ф-лы (8.14) имеем

$$C = (f_m \log_s M)(2 \log_2 s) = B \log_2 s^2. \tag{8.16}$$

Но для s-ичной КИМ отношение мощности сигнала к мощности шума на входе определяется ф-лой (7.81), из которой

$$s^2 = 1 + \frac{12}{K^2} \frac{S_i}{N_i} .$$

Следовательно,

$$C = B \log_2 \left(1 + \frac{12}{K^2} \frac{S_i}{N_i}\right) . \qquad (8.17)$$

Формула (8.17) определяет емкость канала, теоретически необходимую для передачи квантованного сообщения $f(t)$. В действительности, однако, это сообщение передается по каналу с полосой частот B [ур-ние (8.15)] и мощностями сигнала и шума S_i и N_i соответственно. Следовательно, емкость реально используемого канала согласно ф-ле (8.7).

$$C' = B \log_2 \left(1 + \frac{S_i}{N_i}\right), \text{ дв. ед./сек.} \qquad (8.18)$$

Очевидно, S_i/N_i в реальной системе КИМ в $K^2/12$ раз превышает значение, требуемое для идеальной системы. В гл. 9 будет показано, что приемлемой величиной K для сравнительно малой вероятности ошибки (10^{-6}) является $K=10$. Следовательно, требуемая для КИМ мощность приблизительно в 100/12 раз (9,2 *дб*) превышает теоретически необходимую для идеальной системы.

Для вероятности ошибки 10^{-5} различие составляет около 8 *дб*. Характеристики идеальной системы и КИМ для различных значений s показаны на рис. 8.3. Средняя вероятность ошибки при приеме составляет 10^{-5}. Реальная кривая смещена вправо на 8 *дб* относительно идеальной.

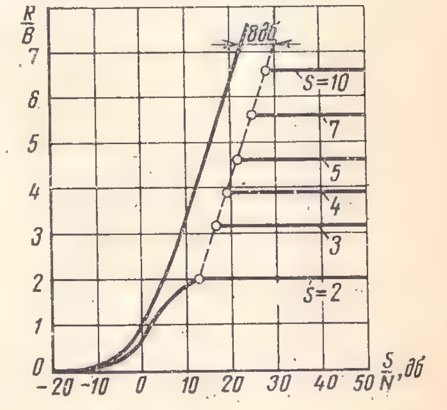


Рис. 8.3. К сравнению эффективности идеальной системы и КИМ

На рис. 8.3 наблюдается эффект «насыщения» для максимальной скорости передачи при заданном значении s. Так, при $s=2$ (двоичная КИМ) скорость передачи на единицу полосы частот не превышает 2 *дв. ед.* Причину этого нетрудно понять. Двоичный импульс может переносить максимум 1 *дв. ед.* информации. По системе с полосой частот B, *гц*, можно передавать $2B$ импульсов в секунду. Итак, скорость передачи $R = 2B$ *дв. ед/сек* и

$$R/B = 2 \text{ дв. ед./сек} \cdot \text{гц}. \qquad (8.19)$$

Если отношение сигнал/шум мало́, импульсы не могут быть уверенно различимы при приеме и, следовательно, с большой вероятностью возникают ошибки. Это приводит к тому, что при ма-

лом отношении сигнал/шум величина R/B оказывается меньше двух. По мере улучшения отношения сигнал/шум вероятность ошибки уменьшается и R/B достигает своего предельного значения, равного двум. При определенном значении сигнал/шум импульсы начинают отчетливо различаться в шуме и дальнейшее увеличение мощности сигнала (или отношения сигнал/шум) не влияет на характеристики системы. Это и является причиной «насыщения». Уровень «насыщения» может быть увеличен путем увеличения s, как показано на рис. 8.3. В общем случае максимум переданной информации с помощью s-ичного импульса равен $\log_2 s$, *дв. ед./сек.* По системе с полосой частот B можно передавать 2 B импульсов в секунду. Следовательно, максимальная скорость передачи информации

$$R = 2B \log_2 s, \text{ дв. ед./сек}$$

и

$$R/B = 2 \log_2 s, \text{ дв. ед./сек·гц.} \quad (8.20)$$

Так, для $s=3$

$$(R/B)_{\text{макс}} = 2 \log_2 3 = 3{,}16 \text{ дв. ед./сек·гц,}$$

для $s=4$

$$R/B = 2 \log_2 4 = 4 \text{ дв. ед./сек·гц.}$$

Заметим, что емкость канала, определяемая формулой Хартли—Шеннона не может быть достигнута при конечном кодировании. Она может быть реализована при блочном кодировании, когда последовательность N символов рассматривается как единственный символ, и в пределе $N \to \infty$. Сложность такого кодирования столь велика, что на практике довольствуются решениями, несколько уступающими теоретическим.

ЗАДАЧИ

8.1. Колебание $Sa\ (2000\ \pi t)$, *в,* должно быть передано с неопределенностью, не превышающей 1/80 *в.* Определите необходимую емкость канала (см. рис. 1.12).

8.2. Повторите задачу 8.1 для колебания $[Sa(2000\ \pi t)]^2$, если неопределенность не превышает 1/64 *в* (см. рис. 1.12).

8.3. Составьте схему передачи непрерывного сигнала $f(t)$, показанного на рис. 8.1 с неопределенностью, не превышающей 0,1 *в,* используя канал с полосой частот $(f_m/2)$ *гц.* Предполагается, что спектр сигнала $f(t)$ ограничен частотой f_m *гц.*

8.4. Повторите задачу 8.3, если неопределенность не превышает 0,025 *в.*

8.5. В примере 8.1 (стр. 292) было найдено, что количество информации на кадр телевизионного изображения составляет около $9{,}96 \times 10^6$ *дв. ед.* Диктор радиовещания пытается устно описать телевизионное изображение, используя 1000 слов из словаря в 10 000 слов. Предполагается, что каждое из 10 000 слов словарного запаса диктора встречается с одинаковой вероятностью при описании этого изображения. Определите количество информации, переданное диктором при описании изображения. Считаете ли вы, что диктор мог бы дать детальное описание изображения, используя 1000 слов? Означает ли пословица «лучше один раз увидеть, чем 1000 раз услышать» преувеличение или принижение действительности?

8.6. При фототелеграфной передаче изображения кадр состоит из $2{,}5 \times 10^6$ элементов. Для хорошего воспроизведения необходимы 12 градаций (уровней) яркости. Предполагается, что все уровни яркости встречаются с одинаковой вероятностью. Определите полосу частот канала, необходимую для передачи одного изображения каждые три минуты. Предполагается, что отношение мощности сигнала к мощности шума в канале составляет 30 *дб* (1000).

8.7. Рассмотрим передачу синоптиком сведений о погоде по телеграфу. Имеются четыре возможных сообщения: солнечно, облачно, дождливо, туман. Если каждое из сообщений равновероятно, то каково минимальное число импульсов, необходимых для передачи сообщения? Составьте типовое кодовое представление для четырех сообщений при использовании двоичных импульсов.

Предположите теперь, что вероятности четырех сообщений равны 1/4, 1/8, 1/8 и 1/2 соответственно. В этом случае информация, приходящаяся в среднем на одно сообщение, меньше 2 *дв. ед.* Это видно из ф-лы (8.2а).

В данном случае P_1, P_2, P_3 и P_4 равны 1/4, 1/8, 1/8 и 1/2. Поэтому можно использовать код, который потребует в среднем меньше 2 *дв. ед.* на сообщение. Докажите, что применение следующего кода действительно потребует только 1,75 *дв. ед./сообщение*

солнечно	10	дождливо	110
облачно	110	туман	0

Заметим, что этот код декодируется однозначно, т. е. любая возможная последовательность, образованная этим кодом, является однозначно декодируемой. Читатель может проверить, что не существует другого однозначно декодируемого кода, который был бы лучше данного. Это объясняется тем, что средняя информация на сообщение в данном случае (см. ф-лу 8.2а) равна 1,75 *дв. ед.*

8.8. В сумке имеется 81 монета; 80 монет совершенно одинаковы, последняя же монета несколько тяжелее остальных. Какое количество информации потребуется, чтобы найти эту монету?

Указание. Определите минимальное число взвешиваний, необходимое для отыскания утяжеленной монеты.

8.9. Пусть известно, что одна из монет, о которых шла речь в задаче 8.8, отличается по весу (но неизвестно, является ли она тяжелее или легче других). Какое количество информации потребуется, чтобы найти эту монету и определить, легче ли она или тяжелее других?

ГЛАВА

9

Элементы цифровой связи

9.1. Введение

В предыдущих главах основное внимание уделялось передаче непрерывных сигналов. Существует бесконечное множество возможных колебаний, которые могут быть образованы непрерывными сигналами. С другой стороны, возможны случаи, когда интерес представляет передача одного из конечного числа колебаний или сообщений. Простым примером такого случая служит передача текста посредством некоторого кода (например, кода Морзе). Код Морзе включает в себя 27 символов или сообщений (26 букв и интервал). Эти символы передаются различными комбинациями посылок и пауз. Следовательно, задача связи сводится к передаче последовательности колебаний, каждое из которых выбирается из полностью определенного и конечного множества. Такой тип связи называется *цифровой связью*, в отличие от передачи непрерывных колебаний, как, например, в системах радио или телевизионного вещания, где множество возможных колебаний бесконечно.

Теперь очевидно основное различие между системами передачи цифровых (или дискретных) и непрерывных (или аналоговых) сообщений. В цифровых системах под связью понимается передача и прием одного из конечного множества известных колебаний, тогда как при непрерывной связи имеется бесчисленное множество сообщений, а соответствующие колебания полностью неизвестны. Следует отметить, что рассмотренная в предыдущей главе система КИМ есть способ цифровой связи, используемый для передачи непрерывных сообщений. Такое преобразование стало возможным благодаря процессу квантования. Этот процесс в действительности аппроксимирует непрерывные сообщения так, чтобы их можно было представить лишь определенными дискретными уровнями. По существу, это цифровая форма непрерывного сообщения. Сообщения теперь могут быть переданы посредством конечного числа символов или уровней.

В цифровых системах задача приема несколько проще, чем в непрерывных. Во время передачи колебания искажаются шумом в канале. Когда такой сигнал поступает к приемнику, необходимо принять решение о том, какое из n возможных известных колебаний действительно было передано. Как только такое решение при-

нимается, переданное колебание воспроизводится совершенно точно, без какого-либо шума. В этом смысле шум канала не искажает форму сигнала. Однако он приводит к тому, что некоторые решения являются ошибочными, причем число ошибок увеличивается с ростом шума в канале. Следовательно, наиболее подходящим критерием обнаружения сигналов в цифровых системах связи является вероятность ошибки. В этой главе рассматривается задача оптимального приема. Ограничимся двоичными системами связи, т. е. системами, в которых используются только два символа. Такой вид связи наиболее часто встречается на практике.

9.2. Прием двоичных сигналов. Согласованный фильтр

В двоичных системах связи сообщения передаются посредством двух символов (так же, как при двоичной КИМ). Один символ представляется импульсом $s(t)$ (рис. 9.1*а*), другой — отсутствием импульса (паузой)[1]. Пусть длительности импульса $s(t)$ и

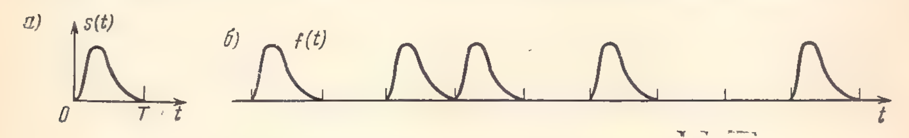


Рис. 9.1. К пояснению двоичной системы связи:
а) импульс $s(t)$; *б)* передаваемый двоичный сигнал

паузы одинаковы и равны T *сек*. Следует снова подчеркнуть, что при обнаружении импульса его форма не имеет значения, так как она заранее известна. Нужно определить лишь, присутствует импульс $s(t)$ или нет. Детектор приемника должен быть поэтому решающим устройством. Он должен анализировать приходящий сигнал каждые T *сек* и выносить решение о наличии или отсутствии импульса. Оптимальным приемником будет тот, который выносит решения с минимальной вероятностью ошибки.

Решение можно сделать более достоверным, если предварительно пропустить сигнал через фильтр, который увеличивает полезный сигнал и в то же время ослабляет шум. Найдем фильтр, на выходе которого в некоторый момент составляющая сигнала максимальна, а составляющая шума ослаблена. Этим будет обеспечено резкое различие между сигналом и шумом на выходе фильтра, и если импульс $s(t)$ присутствует, то на выходе фильтра в этот момент появится большой пик. Если же импульс отсутствует, то пик на выходе не появится. Иными словами, фильтр должен максимизировать отношение амплитуды сигнала к амплитуде шу-

---

[1]) Такой способ называется передачей с пассивной паузой. Возможна передача с активной паузой, т. е. при помощи двух различных сигналов $s_1(t)$ и $s_2(t)$. — *Прим. ред.*

ма на выходе в некоторый момент времени. При доказательстве удобнее оперировать с квадратами амплитуд, поэтому будем искать фильтр, обеспечивающий наибольшее отношение квадрата амплитуды сигнала к квадрату амплитуды шума.

Пусть входной сигнал есть $s(t)+n(t)$, где $s(t)$ — полезный сигнал; $n(t)$ — шум в канале; $s_o(t)+n_o(t)$ — колебание на выходе фильтра (рис. 9.2). Составляющая сигнала на выходе равна $s_o(t)$, составляющая шума — $n_o(t)$. Итак, необходимо максимизировать отношение $s_o^2(t)/n_o^2(t)$ в некоторый момент $t=t_m$ (момент вынесения решения). Заметим, что $n_o(t)$ — случайный сигнал и, следовательно, не может быть определен точно. Поэтому следует рассматривать среднее значение квадрата $\overline{n_o^2(t)}$.

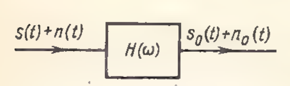


Рис. 9.2. Линейный фильтр с передаточной функцией $H(\omega)$

Другими словами, будем максимизировать отношение ρ, определяемое выражением:

$$\rho=s_o^2(t_m)/\overline{n_o^2(t_m)}. \qquad (9.1)$$

Пусть $S(\omega)$ есть преобразование Фурье $s(t)$ и $H(\omega)$ передаточная функция искомого оптимального фильтра. Тогда

$$s_o(t)=\mathcal{F}^{-1}[S(\omega)H(\omega)]=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(\omega)S(\omega)e^{i\omega t}d\omega$$

и

$$s_0(t_m)=\frac{1}{2\pi}\int_{-\infty}^{\infty}H(\omega)S(\omega)e^{i\omega t_m}d\omega. \qquad (9.2)$$

Средний квадрат составляющей шума можно выразить через спектр плотности мощности шума на выходе фильтра [ф-ла (2.37)]. Если $S_n(\omega)$ — спектр плотности мощности шума $n(t)$ на входе фильтра, то $|H(\omega)|^2S_n(\omega)$ — спектр плотности мощности шума на выходе фильтра. Следовательно,

$$\overline{n_o^2(t)}=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_n(\omega)|H(\omega)|^2d\omega. \qquad (9.3)$$

Заметим, что среднее значение квадрата $n_o(t)$ не зависит от t. Следовательно,

$$\overline{n_o^2(t_m)}=\frac{1}{2\pi}\int_{-\infty}^{\infty}S_n(\omega)|H(\omega)|^2d\omega. \qquad (9.4)$$

Пусть шум $n(t)$ в канале является белым со спектральной плотностью мощности $N/2 : S_n(\omega) = N/2$. Тогда

$$\overline{n_o^2(t)} = \frac{N}{4\pi} \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega. \qquad (9.5)$$

Подставив выражения (9.2) и (9.4) в ф-лу (9.1), получим

$$\rho = s_o^2(t_m)/\overline{n_o^2(t_m)} = \left| \int_{-\infty}^{\infty} H(\omega) S(\omega) e^{i\omega t_m} d\omega \right|^2 / \pi N \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega. \qquad (9.6)$$

Заметим, что поскольку $s_o(t)$ — действительная функция, то $s_o^2(t) = |s_o(t)|^2$.

Далее используем неравенство Шварца, доказательство которого приводится ниже. Одна из форм этого неравенства устанавливает, что если $F_1(\omega)$ и $F_2(\omega)$ — комплексные функции, то

$$\left| \int_{-\infty}^{\infty} F_1(\omega) F_2(\omega) d\omega \right|^2 \leqslant \int_{-\infty}^{\infty} |F_1(\omega)|^2 d\omega \int_{-\infty}^{\infty} |F_2(\omega)|^2 d\omega. \qquad (9.7а)$$

Равенство имеет место, только если

$$F_1(\omega) = k F_2^*(\omega), \qquad (9.7б)$$

где k — произвольная постоянная.

Если обозначить $F_1(\omega) = H(\omega)$, $F_2(\omega) = S(\omega) e^{i\omega t_m}$, то

$$\left| \int_{-\infty}^{\infty} H(\omega) S(\omega) e^{i\omega t_m} d\omega \right|^2 \leqslant \int_{-\infty}^{\infty} |H(\omega)|^2 d\omega \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega. \qquad (9.8)$$

Подстановка неравенства (9.8) в ф-лу (9.6) дает

$$\frac{s_o^2(t_m)}{\overline{n_o^2(t_m)}} \leqslant \frac{1}{\pi N} \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega.$$

Следовательно,

$$\rho_{\text{макс}} = \left. \frac{s_o^2(t_m)}{\overline{n_o^2(t_m)}} \right|_{\text{макс}} = \frac{1}{\pi N} \int_{-\infty}^{\infty} |S(\omega)|^2 d\omega. \qquad (9.9)$$

Это выражение справедливо в том случае, если неравенство (9.8) обращается в равенство, что возможно только при условии

$$H(\omega) = k S^*(\omega) e^{-i\omega t_m} = k S(-\omega) e^{-i\omega t_m}, \qquad (9.10)$$

где k — произвольная постоянная.

Импульсная реакция $h(t)$ оптимального фильтра

$$h(t) = \mathcal{F}^{-1}[H(\omega)] = \mathcal{F}^{-1}\left[k S(-\omega) e^{-i\omega t_m}\right].$$

Заметим, что обратное преобразование Фурье $S(-\omega)$ равно $s(-t)$, а $e^{-i\omega t_m}$ представляет собой временной сдвиг на t_m *сек.*

Следовательно,

$$h(t) = k s(t_m - t). \tag{9.11}$$

Для удобства положим $k=1$.

Как уже упоминалось, полезный сигнал $s(t)$ имеет конечную длительность. Пусть сигнал $s(t)$ равен нулю за пределами интервала $(0, T)$, как показано на рис. 9.3*а*. Сигнал $s(t_m - t)$ может быть получен поворотом $s(t)$ относительно вертикальной оси и сдвигом вправо на t_m *сек*. На рис. 9.3 показаны три случая: $t_m < T$, $t_m = T$ и $t_m > T$. Видно, что импульсная реакция $h(t)$ при $t_m < T$ соответствует физически нереализуемому фильтру (рис. 9.3*в*). Для выполнения условия физической реализуемости необходимо, чтобы $t_m \geqslant T$, как показано на рис. 9.3*г* и *д*. Оба этих фильтра будут давать нужный результат. Однако желательно иметь t_m как можно меньше, чтобы быстрее вынести решение. Следовательно, $t_m = T$ предпочтительней, чем $t_m > T$. Отсюда импульсная реакция оптимального фильтра представляет собой зеркальное отображение принимаемого сигнала $s(t)$ относительно вертикальной оси, сдвинутое вправо на T *сек*. Этот приемник называется *согласованным фильтром* или *согласованным приемником*.

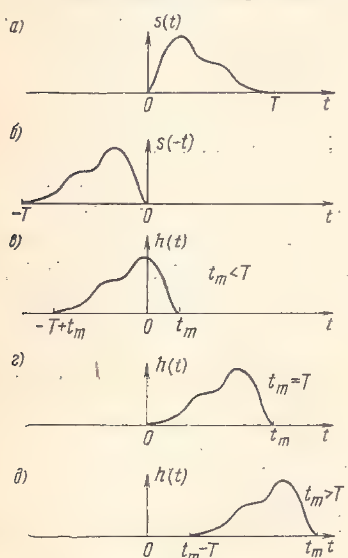


Рис. 9.3. К пояснению импульсной реакции согласованного фильтра:

а) полезный сигнал; *б)* перевернутый сигнал; *в)* импульсная реакция физически нереализуемого фильтра; *г, д)* то же, физически реализуемого фильтра

В согласованном фильтре отношение сигнал/шум (9.1) становится максимальным в момент t_m, который является также моментом, когда весь сигнал $s(t)$ «войдет» в фильтр (рис. 9.4). Следует отметить, что согласованный фильтр является оптимальным среди всех линейных фильтров. Вообще, если снять ограничение на линейность фильтра, можно получить еще большее отношение сигнал/шум. Конечно, фильтр тогда будет нелинейным.

Максимальное отношение сигнал/шум, достигаемое согласованным фильтром, определяется по ф-ле (9.9). Заметим, что энергия сигнала

$$E = \int_{-\infty}^{\infty} s^2(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |S(\omega)|^2\,d\omega.$$

Следовательно,

$$\rho = \frac{\overline{s_o^2(t_m)}}{\overline{n_o^2(t_m)}} = \frac{E}{N/2} = \frac{2E}{N} =$$

$$= \frac{\text{энергия сигнала } s(t)}{\text{спектральная плотность мощности шума на входе}}. \qquad (9.12)$$

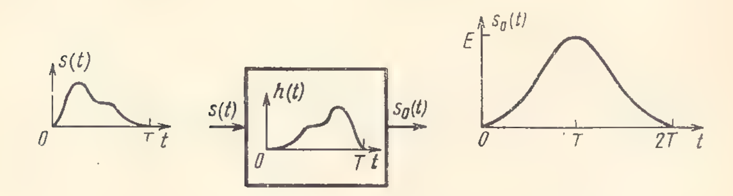


Рис. 9.4. Сигналы на входе и выходе согласованного фильтра

Пиковое значение сигнала $s_o(t_m)$ получается подстановкой ф-лы (9.10) в (9.2):

$$s_o(t_m) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} |S(\omega)|^2 d\omega = E. \qquad (9.13)$$

Таким образом, пиковое значение составляющей сигнала на выходе фильтра появляется в момент $t = t_m$ и численно равно энергии сигнала $s(t)$ (см. рис. 9.4). Это замечательный результат. Пиковое значение составляющей сигнала не зависит от формы колебания $s(t)$, а зависит только от его энергии. Среднее значение квадрата шума на выходе фильтра получается подстановкой ф-лы (9.13) в (9.12)

$$\overline{n_o^2(t_m)} = NE/2. \qquad (9.14)$$

Рис. 9.5. Реализация согласованного фильтра в форме коррелятора

Другая форма реализации согласованного фильтра показана на рис. 9.5. Если на входе согласованного фильтра действует сигнал $f(t)$, то выходной сигнал $r(t) = \int\limits_{-\infty}^{\infty} f(x)h(t-x)dx$, где $h(t) = s(t_m - t)$ — импульсная реакция фильтра. Следовательно, $r(t) = \int\limits_{-\infty}^{\infty} f(x)s(x + t_m - t)dx$.

В момент вынесения решения $t=t_m$ имеем

$$r(t_m)=\int_{-\infty}^{\infty} f(x)\,s(x)\,dx. \qquad (9.15)$$

Значение $r(t_m)$ можно получить с помощью устройства (известного как временной коррелятор), показанного на рис. 9.5. В этом устройстве принимаемый сигнал $s(t)+n(t)$ перемножается с сигналом $s(t)$. Это не что иное, как синхронное детектирование. Таким образом, прием с помощью согласованного фильтра является, по существу, синхронным приемом.

Докажем неравенство Шварца, которое устанавливает, что если $F_1(\omega)$ и $F_2(\omega)$ являются комплексными функциями, то

$$\left|\int_{-\infty}^{\infty} F_1(\omega)\,F_2(\omega)\,d\omega\right|^2 \leqslant \left(\int_{-\infty}^{\infty}|F_1(\omega)|^2\,d\omega\right)\left(\int_{-\infty}^{\infty}|F_2(\omega)|^2\,d\omega\right).$$

Пусть $\Phi(\omega)=F_2^*(\omega)/[\int_{-\infty}^{\infty}|F_2(\omega)|^2 d\omega]^{1/2}$ и $\alpha=\int_{-\infty}^{\infty}F_1(\omega)\Phi^*(\omega)d\omega$. Тогда, поскольку

$$[F_1(\omega)-\alpha\,\Phi(\omega)]\left[F_1^*(\omega)-\alpha^*\,\Phi^*(\omega)\right]=|F_1(\omega)-\alpha\,\Phi(\omega)|^2\geqslant 0,$$

имеем

$$\int_{-\infty}^{\infty}|F_1|^2\,d\omega+|\alpha|^2\int_{-\infty}^{\infty}|\Phi|^2\,d\omega-\alpha\int_{-\infty}^{\infty}\Phi\,F_1^*\,d\omega-\alpha^*\int_{-\infty}^{\infty}\Phi^*\,F_1\,d\omega>0.$$

Но из формулы для $\Phi(\omega)$ следует, что $\int_{-\infty}^{\infty}|\Phi|^2 d\omega=1$, а из формулы для α, что $\int_{-\infty}^{\infty}\Phi F_1^*\,d\omega=\alpha^*$. Таким образом, полученное выражение можно переписать в виде

$$\int_{-\infty}^{\infty}|F_1|^2\,d\omega+|\alpha|^2-\alpha\alpha^*-\alpha^*\alpha>0 \qquad \text{или} \qquad \int_{-\infty}^{\infty}|F_1(\omega)|^2\,d\omega-|\alpha|^2\geqslant 0.$$

Подстановка формул для $\Phi(\omega)$ и α в последнее выражение дает

$$\int_{-\infty}^{\infty}|F_1(\omega)|^2\,d\omega\geqslant\left|\int_{-\infty}^{\infty}F_1(\omega)\,F_2(\omega)\,d\omega\right|^2\Big/\int_{-\infty}^{\infty}|F_2(\omega)|^2\,d\omega,$$

что и требовалось доказать.

Заметим, что предпоследнее неравенство обращается в равенство, если и только если $\int_{-\infty}^{\infty}|F_1(\omega)|^2 d\omega=|\alpha|^2$. Из выражения для $\Phi(\omega)$ следует, что это возможно только при $F_1(\omega)=kF_2^*(\omega)$, где k — произвольная постоянная.

9.3. Оптимизация обнаружения сигнала согласованным фильтром

ОПТИМАЛЬНЫЙ ПОРОГ РЕШЕНИЯ

Согласованный фильтр максимизирует отношение сигнал/шум в момент $t_m (t_m = T)$. Поэтому вопрос о том, присутствует ли сигнал $s(t)$, решается путем наблюдения выходного колебания $r(t)$ в момент $t=T$: $r(t)=s_o(T)+n_o(T)$. Подставив ф-лу (9.13) в последнее выражение, получим

$$r(T)=E+n_0(T), \tag{9.16}$$

где $n_o(T)$ — случайное колебание.

Следовательно, если сигнал $s(t)$ присутствует на входе фильтра, то $r(T)$ представляет собой сумму постоянной величины E и случайной величины $n_o(T)$. Если же сигнал $s(t)$ отсутствует, то на выходе фильтра будет только шумовая составляющая

$$r(T)=n_o(T). \tag{9.17}$$

Случайный характер составляющей $n_o(T)$ приводит к ошибочным решениям. Возможно, что сигнал $s(t)$ присутствует, но составляющая $n_o(T)$ имеет большое отрицательное значение. Тогда $r(T)$ окажется очень малым. С другой стороны, даже если $s(t)$ отсутствует, $n_o(t)$ может принять большое положительное значение. Это приведет к тому, что $r(T)$ будет очень большим. Таким образом, нет надежного способа вынести решение о присутствии или отсутствии $s(t)$. Очевидно, однако, что когда $r(T)$ велико, то более вероятно, что $s(t)$ присутствует. Если же $r(T)$ очень мало, то более вероятно, что сигнал отсутствует. Независимо от того, какое выносится решение, всегда имеется некоторая вероятность ошибки. Поэтому следует найти такое правило решения, которое минимизирует вероятность ошибки.

Пусть правило решения заключается в следующем: «сигнал есть», если $r(T)>a$ и «сигнала нет», если $r(T)<a$. Найдем оптимальный порог решения a, при котором вероятность ошибок минимальна. Для этого рассмотрим более детально случайный характер составляющей $n_o(T)$. Случайные значения $n_o(T)$ подчиняются определенному распределению, чаще всего гауссову[1]. Это означа-

---

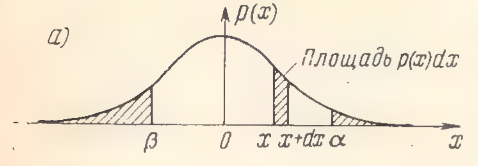
[1]) Используя центральную предельную теорему теории вероятностей, можно показать, что распределение вероятностей сложного сигнала, образованного из большого числа относительно независимых сигналов, стремится к гауссову. Большинство шумовых сигналов является результатом сравнительно многочисленных независимых возмущений. Дробовой и тепловой шумы попадают в эту категорию. Поэтому предположение о гауссовом шуме справедливо в большинстве случаев.

ет, что относительная частота появления значений шума описывается гауссовой кривой. Распределение $p(x)$ называется функцией плотности вероятности амплитуды x и записывается в виде

$$p(x)=\frac{1}{\sigma_x\sqrt{2\pi}}e^{-x^2/\left(2\sigma_x^2\right)}, \tag{9.18}$$

где σ_x^2 — среднее значение квадрата шума.

Это распределение показано на рис. 9.6*а*, из которого видно, что распределение значений шума симметрично относительно



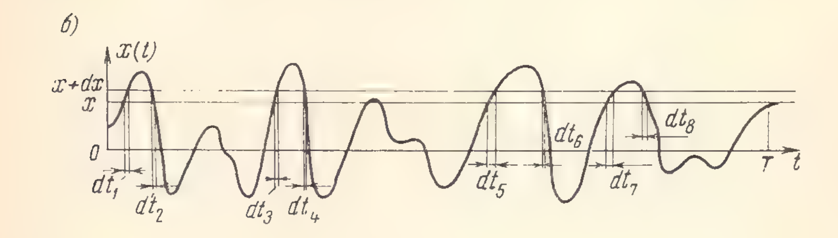


Рис. 9.6. К определению функции плотности вероятности шума: *а*) функция плотности вероятности $p(x)$; *б*) наблюдаемая реализация шума

$x=0$. Шум с одинаковой вероятностью принимает положительные и отрицательные значения, следовательно, он имеет нулевое среднее значение. Плотность вероятности представляет собой относительную частоту появления заданного значения шума. Эта функция так нормирована, что площадь $p(x)dx$ (показанная на рис. 9.6*а*) представляет собой вероятность того, что значения шума находятся в интервале $(x, x+dx)$. Таким образом, если наблюдать шум на интервале 0, T $(T\to\infty)$, как показано на рис. 9.6*б*, то можно заметить, что амплитуда находится в интервале $(x, x+dx)$ в течение некоторого времени dT, равного $dT=\sum_{i=1}^{\infty}dt_i$. Следовательно, относительная частота наблюдения шума в интервале $(x, x+$ $+dx)$ равна dT/T. Поэтому $p(x)dx=\Sigma\, dt_i/T$.

Очевидно, вероятность наблюдения x в интервале (x_1, x_2) определяется площадью под кривой $p(x)$ в интервале (x_1, x_2)

$$P(x_1<x<x_2)=\int_{x_1}^{x_2}p(x)\,dx. \tag{9.19}$$

Аналогично, вероятность наблюдения $x>\alpha$

$$P(x>\alpha)=\int\limits_{\alpha}^{\infty}p(x)\,dx \qquad (9.20)$$

и

$$P(x<|\beta|)=\int\limits_{-\infty}^{\beta}p(x)\,dx. \qquad (9.21)$$

Среднее значение квадрата шума на выходе согласованного фильтра равно $NE/2$ [ф-ла (9.14)]. Следовательно, в ф-ле (9.18) $\sigma_x^2=NE/2$ и

$$p(x)=\frac{1}{\sqrt{\pi NE}}\,e^{-x^2/NE}. \qquad (9.22)$$

Если на входе фильтра сигнал $s(t)$ отсутствует, на его выходе наблюдается составляющая $n_o(T)$, которая имеет распределение

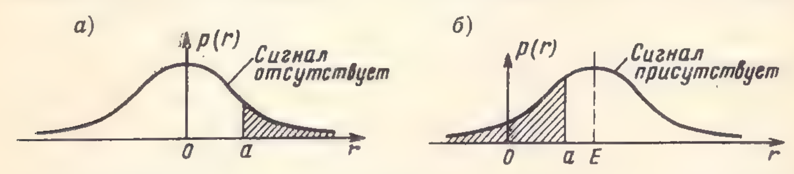


Рис. 9.7. Функции плотности вероятности колебания на выходе согласованного фильтра в момент решения

вероятностей (9.22). Оно показано на рис. 9.7*а*. Если обозначить выходное колебание через r, то $r=n_o(T)$ и

$$p(r)=\frac{1}{\sqrt{\pi NE}}\,e^{-r^2/NE}. \qquad (9.23)$$

Если сигнал $s(t)$ присутствует, то $r=E+n_o(T)$ и распределение вероятностей смещается на величину постоянной E:

$$p(r)=\frac{1}{\sqrt{\pi NE}}\,e^{-(r-E)^2/NE}. \qquad (9.24)$$

Это распределение показано на рис. 9.7*б*. Оба распределения вместе приведены на рис. 9.8.

В соответствии с принятым выше правилом решение «сигнал есть» выносится, если $r>a$ и ««сигнала нет» — если $r<a$. Из рис. 9.7*а* видно, что в некоторые моменты времени $r>a$, хотя сигнал отсутствует. Вероятность $r>a$ в отсутствие сигнала равна заштрихованной площади на рис. 9.7*а*. Такое решение приводит к ошибке (называемой *ложной тревогой*).

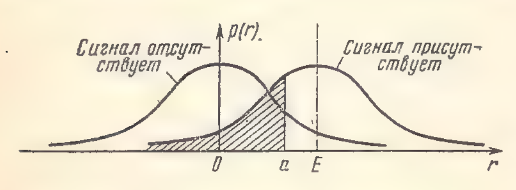


Рис. 9.8. К определению оптимального порога решения. Средняя вероятность ошибки минимальна при $a=E/2$

С другой стороны, если в присутствие сигнала выходное колебание окажется ниже порога a, то будет принято решение «сигнала нет», хотя на самом деле он присутствует. Этот тип ошибки называется **пропуском сигнала**, а ее вероятность определяется заштрихованной площадью на рис. 9.7*б*.

Если сигнал $s(t)$ с одинаковой вероятностью присутствует и отсутствует, то это означает, что в среднем половину времени сигналы $s(t)$ передаются, а половину — не передаются. Вероятность ошибочных решений минимальна, если минимальна сумма двух заштрихованных площадей на рис. 9.8. Как видно из рис. 9.8, это будет, если выбрать

$$a = E/2. \tag{9.25}$$

Эта формула и определяет оптимальное значение порога решения.

ВЕРОЯТНОСТЬ ОШИБКИ

Если присутствие и отсутствие сигнала $s(t)$ равновероятны, то вероятность ошибочных решений равна половине суммарной заштрихованной площади на рис. 9.7*а* и *б*. При оптимальном значении $a=E/2$ обе площади одинаковы. Поэтому вероятность ошибки $P(\varepsilon)$ определяется одной из этих площадей. Из рис. 9.7*а* получаем

$$P(\varepsilon) = \int_a^\infty p(r)\,dr = \frac{1}{\sqrt{\pi N E}} \int_a^\infty e^{-r^2/NE} dr. \tag{9.26}$$

Интеграл в правой части ф-лы (9.26) в замкнутом виде не вычисляется. Он, однако, табулирован в известных стандартных таблицах интеграла вероятности или функции ошибок $\mathrm{erf}(x)$.

Определим функцию ошибок $\mathrm{erf}(x)$, как[1])

$$\mathrm{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy \tag{9.27}$$

и дополнительную функцию ошибок $\mathrm{erf\,c}(x)$ как

$$\mathrm{erf\,c}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-y^2/2}\, dy. \tag{9.28}$$

Из этих определений очевидно, что

$$\mathrm{erf}(x) + \mathrm{erfc}(x) = 1. \tag{9.29}$$

---

[1]) В литературе имеется несколько определений $\mathrm{erf}(x)$ и $\mathrm{erfc}(x)$, которые, по существу, эквивалентны.

При $x>2$ erfc (x) приближенно представляется как

$$\operatorname{erfc}(x) \approx \frac{1}{x\sqrt{2\pi}}\left(1-\frac{1}{x^2}\right)e^{-x^2/2}. \quad (9.30)$$

Ошибка такой аппроксимации составляет — 10% для $x=2$ и меньше 1% для $x>3$.

Используя определение (9.27), можно записать

$$P(\varepsilon)=\operatorname{erfc}\left(\frac{a}{\sqrt{NE/2}}\right). \quad (9.31)$$

Но поскольку $a=E/2$, то

$$P(\varepsilon)=\operatorname{erfc}\left(\sqrt{E/2N}\right). \quad (9.32)$$

Вероятность ошибки $P(\varepsilon)$ как функция E/N представлена на рис. 9.9.

Рис. 9.9. Зависимость вероятности ошибки от отношения энергии сигнала к спектральной плотности мощности шума E/N

Какой смысл имеет вероятность ошибки? Вероятность события означает правдоподобие или относительную частоту этого события. Так, если вынесено N решений ($N\to\infty$), то общее число ошибочных решений N_ε равно $N_\varepsilon=P(\varepsilon)N$. Если $P(\varepsilon)=0{,}01$, то в среднем одно из ста решений будет ошибочным.

Пример 9.1.

В двоичной системе КИМ (гл. 7) сигнал $s(t)$ представляет собой прямоугольный импульс высотой A и длительностью T. Импульсная реакция согласованного фильтра определяется как $h(t)=s(T-t)$. В рассматриваемом случае $h(t)=s(t)$. Фильтр с такой импульсной реакцией может быть реализован в форме устройства, показанного на рис. 9.10. Энергия E сигнала $s(t)$ равна $E=A^2T$, причем $A=K\sigma_n$, где σ_n — эффективное значение шума. Средняя мощность шума равна $\sigma_n^2=N_i=\overline{n^2(t)}$. Поскольку длительность импульса равна T, то в одну секунду передается $1/T$ импульсов. Для их передачи необходима полоса частот B, равная $1/(2T)$. Если спектральная плотность мощности шума равна $N/2$, то

$$N_i=NB=N/(2T)=\sigma_n^2,$$

$$N=2T\sigma_n^2. \quad (9.33)$$

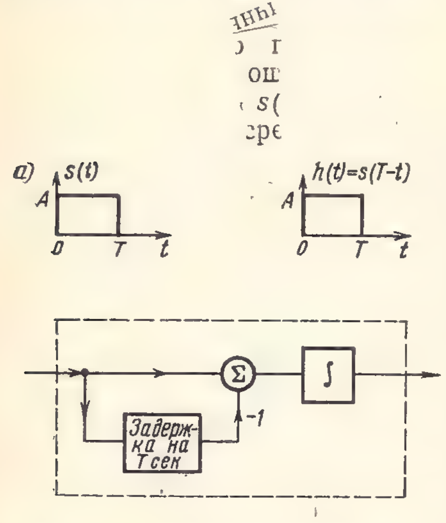


Рис. 9.10. Согласованный фильтр для сигнала прямоугольной формы: *а*) сигнал $s(t)$; *б*) импульсная реакция фильтра $h(t)$; *в*) структурная схема фильтра

Очевидно, $E/N=A^2T/(2T\sigma_n^2)=A^2/(2\sigma_n^2)=K^2\sigma_n^2/(2\sigma_n^2)=K^2/2$. При $K=10$ отношение $E/N=50$ и вероятность ошибки

$$P(\varepsilon) = \operatorname{erf}\sqrt{25} = \operatorname{erf} 5. \tag{9.34}$$

По ф-ле (9.30) получаем

$$P(\varepsilon) \approx 0{,}284 \cdot 10^{-6}. \tag{9.35}$$

Этот результат можно также получить непосредственно из рис. 9.9. При $E/N=50$ $10\lg E/N=16{,}9$ *дб*. Это дает $P(\varepsilon)\approx 0{,}284\cdot 10^{-6}$. Таким образом, если амплитуда импульса в 10 раз превышает эффективное значение шума ($K=10$), то вероятность ошибки имеет порядок 10^{-6}, что приемлемо в большинстве практических случаев.

Здесь предполагалось, что сигнал $s(t)$ имеет форму идеального прямоугольного импульса. Однако вследствие ограничения полосы частот канала импульс в процессе передачи становится трапецеидальным (см. § 2.6). Следовательно, импульсная реакция согласованного фильтра также должна быть трапецеидальной, чтобы обеспечить согласование с формой принимаемого сигнала. Это следует иметь в виду при последующих обсуждениях, где в качестве сигнала $s(t)$ также используются идеализированные импульсы прямоугольной формы.

9.4. Амплитудная манипуляция

Двоичные импульсы КИМ (пример 9.1) легко могут быть переданы по проводам. Но когда передача осуществляется путем излучения в свободном пространстве, следует использовать амплитудномодулированную двоичную КИМ. Амплитудная модуляция переносит низкочастотный спектр двоичной КИМ в область высоких частот (на несущую частоту). Такой вид модуляции называется амплитудной манипуляцией. Один из двоичных импульсов передается синусоидальным импульсом

$$s(t)=\begin{cases} A\sin\omega_c t & \text{при} \quad 0<t<T, \\ 0 & \text{при} \quad t<0,\ t>T. \end{cases}$$

Другой символ передается паузой (сигнал отсутствует). Типичное колебание амплитудной манипуляции представлено на рис. 9.11.

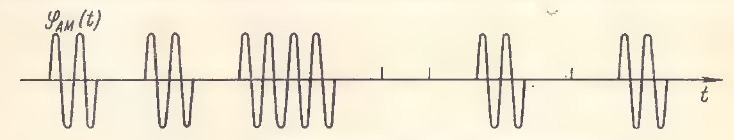


Рис. 9.11. Сигнал амплитудной манипуляции

Найдем теперь оптимальный приемник и вероятность ошибки в случае амплитудной манипуляции. Предполагается, что вероятности присутствия импульса $s(t)$ и его отсутствия одинаковы и равны 0,5.

Импульс $s(t)$ показан на рис. 9.12*а*. Импульсная реакция согласованного фильтра $h(t)=s(T-t)$. Заметим, что $s(T-t)=-s(t)$. Поэтому $h(t)=-s(t)$. Выходной сигнал согласованного фильтра

(когда $s(t)$ присутствует на входе) находится как свертка $s(t)$ и $h(t)$. Результат свертки показан на рис. 9.12*в*. Как и следовало ожидать, выходной сигнал максимален при $t=T$ и численно равен энергии E сигнала $s(t)$:

$$E=A^2T/2. \qquad (9.36)$$

Порог решения равен $E/2$. В данном случае

$$a=A^2T/4. \qquad (9.37)$$

Вероятность ошибки определяется по ф-ле (9.32)

$$P(\varepsilon)=\operatorname{erfc}\left(\sqrt{\frac{E}{2N}}\right)=$$

$$=\operatorname{erfc}\left(\frac{A}{2}\sqrt{\frac{T}{N}}\right). \qquad (9.38)$$

Рис. 9.12. Согласованный фильтр сигнала амплитудной манипуляции:

а) сигнал $s(t)$; *б*) импульсная реакция фильтра $h(t)$; *в*) преобразование сигнала в фильтре; *г*) корреляционный вариант реализации фильтра

Численные значения вероятности ошибки можно непосредственно получить из графика рис. 9.9.

Вероятность ошибки можно выразить через среднюю мощность сигнала. Сигнал $s(t)$ имеет энергию $E=A^2T/2$. Так как сигнал $s(t)$ присутствует в среднем половину времени, то его средняя мощность

$$P_s=\frac{1}{T}\,\frac{E}{2}=\frac{A^2}{4}$$

и

$$P(\varepsilon)=\operatorname{erf}\left(\sqrt{\frac{P_sT}{N}}\right). \qquad (9.39)$$

Согласованный фильтр имеет импульсную реакцию (рис. 9.12*б*)

$$h(t)=-s(t)=\begin{cases}-A\sin\omega_c t & \text{при} \quad 0<t<T,\\ 0 & \text{при} \quad t<0; \quad t>T.\end{cases}$$

В качестве другого решения можно использовать корреляционное устройство, схема которого показана на рис. 9.12*г*. В рассматриваемом случае это не что иное, как синхронный детектор.

9.5. Фазовая манипуляция

В гл. 7 было показано, что более эффективной является двоичная КИМ с биполярными импульсами высотой $A/2$ и $-A/2$. Таким образом, в биполярной КИМ два символа сообщения представля-

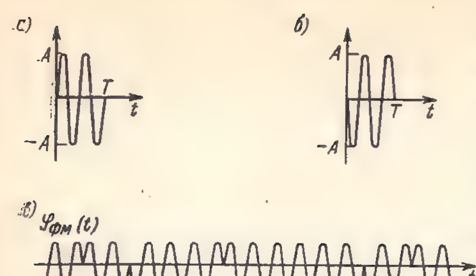
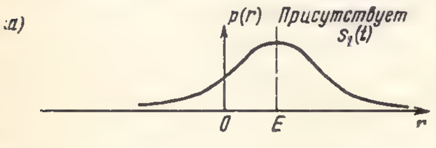
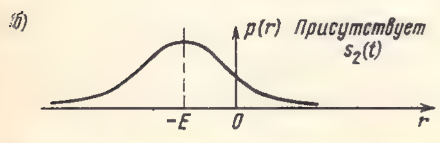


Рис. 9.13. К пояснению фазовой манипуляции:

а) сигнал $s_1(t)$; *б)* сигнал $s_2(t)=-s_1(t)$; *в)* сигнал фазовой манипуляции





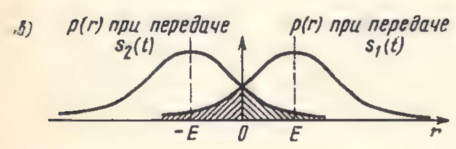


Рис. 9.14. К определению вероятности ошибки при оптимальном приеме сигналов фазовой манипуляции:

а) функция плотности вероятности при передаче $s_1(t)$; *б)* то же, при передаче $s_2(t)$; *в)* совмещенные функции плотности вероятности

ются сигналами $s(t)$ и $-s(t)$. Если в качестве сигналов используются прямоугольные радиоимпульсы (рис. 9.13*а* и *б*), то такой способ передачи называется фазовой манипуляцией. Типичное колебание фазовой манипуляции показано на рис. 9.13*в*. Найдем оптимальный приемник и вероятность ошибки для этого случая.

Два символа передаются посредством колебаний $s_1(t)$ и $s_2(t)=-s_1(t)$. Пусть $s_1(t)=-s_2(t)=s(t)$. Предположим для общности, что колебание $s(t)$ имеет произвольную форму. Единственное ограничение на $s(t)$ состоит в том, что оно имеет длительность T *сек*. Пусть энергия $s(t)$ равна E.

Согласованный фильтр для сигнала $s(t)$ имеет импульсную реакцию $h(t)=s(T-t)$. Если на вход этого фильтра поступает колебание $s_1(t)$, то выходное колебание в момент $t=T$ согласно ф-ле (9.16)

$$r(T)=E+n_o(T). \tag{9.40}$$

Составляющая $n_o(T)$ — случайная переменная, имеющая гауссово распределение со средним значением квадрата $NE/2$. Следовательно,

$$p(r)=\frac{1}{\sqrt{\pi NE}}\,e^{-(r-E)^2/NE}. \tag{9.41}$$

Это распределение показано на рис. 9.14*а*. Аналогично, когда на вход согласованного фильтра поступает сигнал $s_2(t)$, то

$$r(T)=-E+n_o(T). \tag{9.42}$$

Распределение $r(T)$ при передаче сигнала $s_2(t)$ показано на рис. 9.14*б*. Очевидно, в этом случае

$$r(T)=\frac{1}{\sqrt{\pi NE}}\,e^{-(r+E)^2/NE}. \tag{9.43}$$

Оба распределения вместе показаны на рис. 9.14*в*. Определим порог решения, при котором вероятность ошибки минимальна.

Используя те же рассуждения, что и при выводе ф-лы (9.25), видим, что в данном случае оптимальный порог решения

$$a=0. \tag{9.44}$$

Таким образом, если $r(T)>0$, принимается решение «присутствует $s_1(t)$», а если $r(T)<0$ — решение «присутствует $s_2(t)$». Вероятность ошибки определяется площадью под кривой $p(r)$ (на рис. 9.14*б*) в пределах от 0 до ∞.

$$P(\varepsilon)=\frac{1}{\sqrt{\pi N E}}\int_0^\infty e^{-(r+E)^2/NE}\,dr= \tag{9.45}$$

$$=\frac{1}{\sqrt{\pi N E}}\int_E^\infty e^{-x^2/NE}\,dx. \tag{9.46}$$

Этот интеграл в точности совпадает с ф-лой (9.26), за исключением того, что a заменяется на E. Следовательно,

$$P(\varepsilon)=\mathrm{erfc}\left(\sqrt{2EN}\right). \tag{9.47}$$

Для сигналов фазовой манипуляции, показанных на рис. 9.13,

$$P(\varepsilon)=\mathrm{erfc}\left(A\sqrt{T/N}\right). \tag{9.48}$$

Средняя мощность сигнала фазовой манипуляции

$$P_s=A^2/2. \tag{9.49}$$

Следовательно,

$$P(\varepsilon)=\mathrm{erfc}\left(\sqrt{2P_sT/N}\right). \tag{9.50}$$

Сравним последнее выражение с ф-лой (9.39) для амплитудной манипуляции. Из этих двух формул видно, что для достижения заданной вероятности ошибки при амплитудной манипуляции требуется вдвое бо́льшая средняя мощность, чем при фазовой манипуляции. Следовательно, фазовая манипуляция превосходит амплитудную по требуемой средней мощности сигнала на 3*дб*.

9.6. Частотная манипуляция

Частотную манипуляцию можно рассматривать как частотно-модулированную двоичную КИМ. Два символа представляются колебаниями:

$$s_1(t)=\begin{cases}A\sin m\omega_0 t & \text{при}\quad 0<t<T,\\ 0 & \text{при}\quad \text{других } t;\end{cases} \tag{9.51а}$$

$$s_2(t)=\begin{cases}A\sin n\omega_0 t & \text{при}\quad 0<t<T,\\ 0 & \text{при}\quad \text{других } t,\end{cases} \tag{9.51б}$$

где $\omega_0=2\pi/T$.

Сигналы $s_1(t)$ и $s_2(t)$ показаны на рис. 9.15*а*. Это разные сигналы и, следовательно, для их приема требуются два согласованных

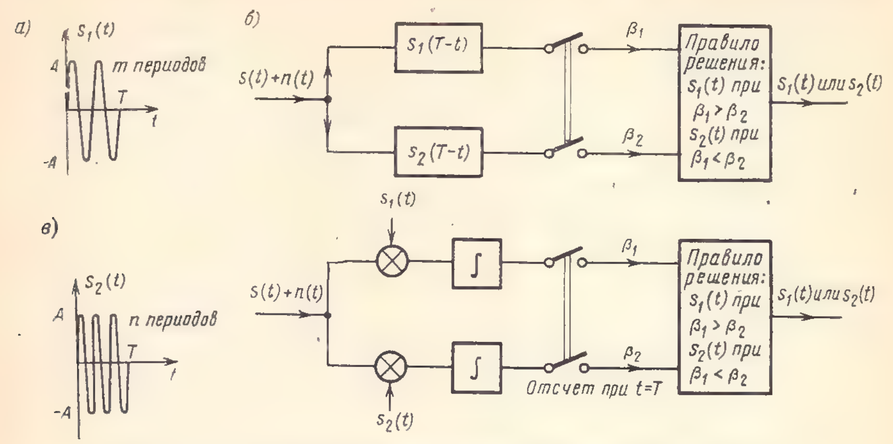


Рис. 9.15. К пояснению оптимального приема сигналов частотной манипуляции:
а) сигналы $s_1(t)$ и $s_2(t)$; *б)* оптимальный приемник с согласованными фильтрами; *в)* то же, с корреляторами

фильтра. Покажем, что оптимальный приемник сигналов частотной манипуляции имеет схемы, приведенные на рис. 9.15*б*, *в*. На рис. 9.15*б* показан приемник с согласованными фильтрами, а на рис. 9.15*в* — с корреляторами.

Входной сигнал фильтра равен $s(t)+n(t)$, где $s(t)$ — либо $s_1(t)$, либо $s_2(t)$. Обозначим входной сигнал через $f(t)=s(t)+n(t)$. Добавим ко входному сигналу — $s_2(t)$, как показано на рис. 9.16*а*. Добавление известного сигнала ко входному сигналу не должно влиять на оптимальный прием[1]). Новый сигнал $f'(t)$ теперь имеет вид

$$f'(t)=s(t)+n(t)-s_2(t).$$

Если сигнал $s(t)$ равен $s_1(t)$, то полезный сигнал в $f'(t)$ будет $s_1(t)-s_2(t)$. Если же сигнал $s(t)$ равен $s_2(t)$, то полезный сигнал в $f'(t)$ будет равен нулю. Таким образом, введение дополнительного колебания позволяет привести частотную манипуляцию к случаю, когда два символа сообщения передаются наличием и отсутствием импульса. Это аналогично случаю двоичных сигналов, рассмотренному в § 9.2. Новый сигнал $s'(t)$ равен теперь $s_1(t)-s_2(t)$. Согласованный фильтр для такого сигнала имеет импульсную реакцию

$$h(t)=s_1(T-t)-s_2(T-t). \tag{9.52}$$

---

[1]) Этот результат является частным случаем теоремы обратимости. Если сигнал подвергается некоторому преднамеренному преобразованию, тогда оптимальная обработка может быть произведена над промежуточным преобразованием, являющимся обратимым. Вычитание $s_2(t)$ есть обратимая операция, поскольку исходный сигнал может быть получен добавлением $s_2(t)$ к $f'(t)$ на рис. 9.16*а*.

Согласованный фильтр показан на рис. 9.16*а*. Энергия E' сигнала $s'(t)$

$$E' = \int_0^T [s_1(t) - s_2(t)]^2\,dt = \int_0^T s_1^2(t)\,dt + \int_0^T s_2^2(t)\,dt - 2\int_0^T s_1(t)\,s_2(t)\,dt, \quad (9.53)$$

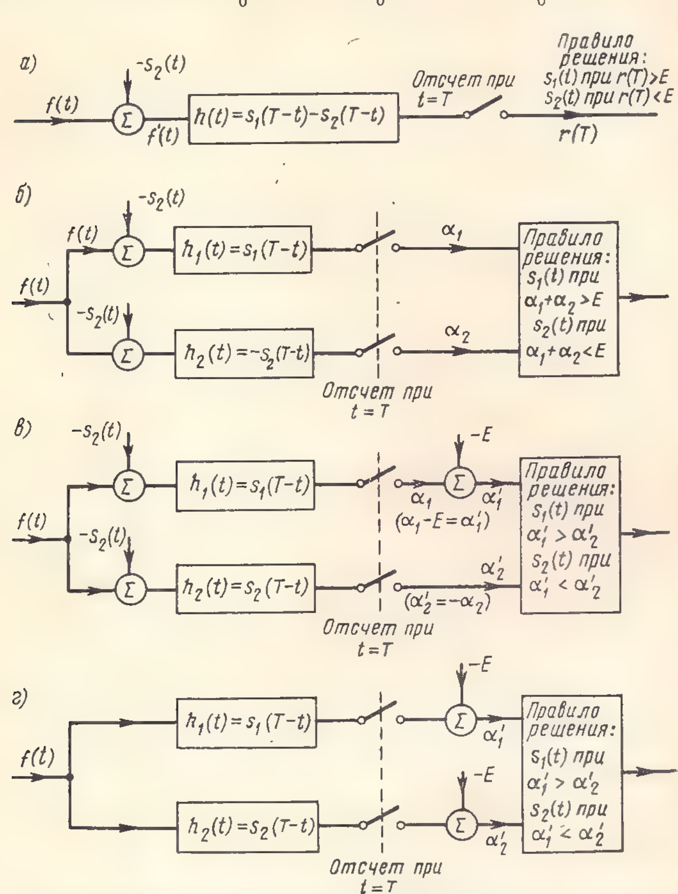


Рис. 9.16. К определению схемы оптимального приемника сигналов частотной манипуляции;
а) оптимальный прием сигнала $f(t)$ с добавлением к нему — $s_2(t)$; *б)* преобразование схемы *(а)* в двухканальный согласованный фильтр; *в)* двухканальный согласованный фильтр со смещенным порогом; *г)* второй вариант двухканального фильтра со смещенным порогом

где $s_1(t)$ и $s_2(t)$ определяются по ф-ле (9.51). Заметим, что последний интеграл в правой части (9.53) равен нулю:

$$\int_0^T \sin m\,\omega_0 t \sin n\,\omega_0 t\,dt = 0 \quad (\omega_0 = 2\pi/T)$$

и
$$\int_0^T s_1^2(t)\,dt=\int_0^T s_2^2(t)\,dt=E=A^2T/2. \qquad (9.54)$$

Следовательно,
$$E'=2E=A^2T. \qquad (9.55)$$

Итак, энергия E' сигнала $s'(t)$ равна $2E(=A^2T)$, где E — энергия сигнала $s_1(t)$ или $s_2(t)$.

Порог решения согласованного фильтра равен $E'/2=E$. Следовательно, решение «присутствует $s_1(t)$» выносится, если $r(T)>E$, и «присутствует $s_2(t)$» — если $r(T)<E$. Такой согласованный фильтр показан на рис. 9.16*а*. Вероятность ошибки для этого фильтра находится по ф-ле (9.32), в которую вместо E следует подставить $E'(=2E)$:
$$(P\,\varepsilon)=\operatorname{erfc}\left(\sqrt{E/N}\right). \qquad (9.56)$$

Средняя мощность сигналов частотной манипуляции
$$P_s=A^2/2=E/T. \qquad (9.57)$$

Следовательно,
$$P(\varepsilon)=\operatorname{erfc}\left(\sqrt{P_s T/N}\right). \qquad (9.58)$$

Это выражение совпадает с ф-лой (9.39) для вероятности ошибки при амплитудной манипуляции. Из изложенного очевидно, что фазовая манипуляция превосходит как амплитудную манипуляцию, так и частотную.

Согласованный фильтр с импульсной реакцией $s_1(T-t)-s_2(T-t)$ может быть образован из двух параллельно включенных согласованных фильтров с импульсными реакциями $s_1(T-t)$ и $s_2(T-t)$, как показано на рис. 9.16*б*. Очевидно, эти два устройства эквивалентны. Выполним дальнейшие преобразования, как показано на рис. 9.16*в*. На рис. 9.16*б* критерием является $\alpha_1+\alpha_2>E$ или $\alpha_1+\alpha_2<E$. Если вычесть E из выходного колебания верхнего фильтра, как показано на рис. 9.16*в*, то критерий решения сведется к $\alpha_1+\alpha_2>0$ или $\alpha_1+\alpha_2<0$. Дополнительно изменим знак импульсной реакции нижнего фильтра. Поэтому на рис. 9.16*в* $\alpha_1'=\alpha_1-E$ и $\alpha_2'=-\alpha_2$. Это приводит к критерию решения $\alpha_1'-\alpha_2'>0$ или $\alpha_1'-\alpha_2'<0$. Таким образом, новое правило решения таково: «присутствует $s_1(t)$», если $\alpha_1'>\alpha_2'$, и «присутствует $s_2(t)$», если $\alpha_1'<\alpha_2'$. Заметим теперь, что выходной сигнал верхнего фильтра в момент $t=T$ при действии на входе сигнала $s_2(t)$ равен нулю:
$$s_2(t)*h_1(t)=\int_0^T A^2\sin n\,\omega_0\tau\sin m\,\omega_0(t-\tau)\,d\tau=0 \quad (\omega_0=2\pi/T).$$

Следовательно, устранение сигнала — $s_2(t)$ со входа верхнего фильтра не влияет на устройство. Далее, сигнал — $s_2(t)$ на входе

нижнего фильтра, согласованного с $s_2(t)$, дает на выходе $-E$ при $t=T$ [см. ур-ние (9.13)]. Поскольку решение выносится в момент $t=T$, то можно устранить $-s_2(t)$ на входе нижнего фильтра и вычесть E из его выходного колебания, как показано на рис. 9.16*г*. Таким образом, устройство на рис. 9.16*г* эквивалентно показанному на рис. 9.16*в*. Отметим, что решение основывается на сравнении двух выходных колебаний схемы рис. 9.16*г*. Поэтому $-E$ к обоим выходам можно не добавлять. В результате приходим к устройству, показанному на рис. 9.15*б*. Соответствующее корреляционное устройство показано на рис. 9.15*в*.

Поскольку все устройства на рис. 9.16 и 9.15*б* эквивалентны, все они имеют одинаковую вероятность ошибки, определяемую по ф-ле (9.56).

9.7. Некоторые замечания о приеме на согласованный фильтр

Выше было показано, что согласованный фильтр эквивалентен временно́му коррелятору (рис. 9.5). В корреляторе входной сигнал $[s(t)+n(t)]$ умножается на $s(t)$. Это, очевидно, синхронный прием (называемый также когерентным).

Ранее упоминалось, что прием посредством согласованного фильтра оптимален в рамках линейных систем. Вообще говоря, можно найти системы лучшего качества, если не ограничиваться классом линейных систем. Можно, однако, показать, что при гауссовом шуме (как это имеет место в большинстве случаев передачи сигналов) согласованный фильтр (корреляционный приемник) является абсолютно оптимальным.

В предыдущих разделах предполагалось, что передатчик и приемник синхронизированы. Для любого данного импульса решение принимается в тот момент, когда импульс полностью «вошел» в

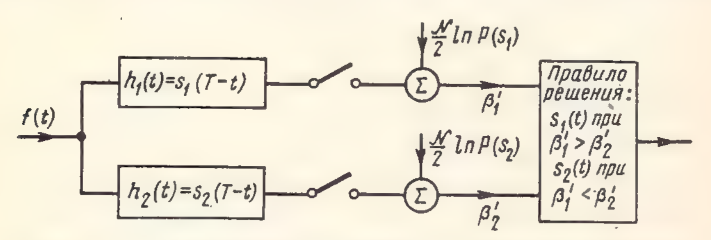


Рис. 9.17. Оптимальный приемник двух неравновероятных сигналов

согласованный фильтр. Таким образом, решения на выходе приемника принимаются каждые T *сек* и моменты отсчета должны быть точно синхронизированы. Кроме того, предполагается, что в момент принятия решения выходной эффект полностью обусловлен действием рассматриваемого импульса и шума. При этом не учитывается возможность взаимной интерференции сигналов, ко-

торая может возникнуть в согласованном фильтре из-за остаточной энергии предыдущего импульса. В идеальном случае остаточный отклик из-за действия $(n-1)$-го импульса равен нулю в момент $t=nT$. Это можно видеть из рис. 9.4. Отклик на импульс длительностью T растягивается на интервал $2T$. Таким образом, отклик на первый импульс будет занимать интервал $(0, 2T)$, тогда как отклик на второй импульс расположится в интервале $(T, 3T)$. Отсчет второго импульса берется в момент $t=2T$, когда отклик от первого импульса равен нулю.

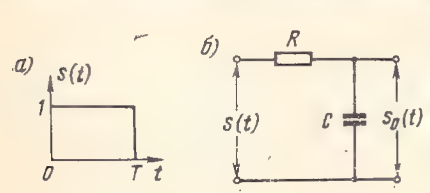
До сих пор предполагалось, что два состояния двоичного символа равновероятны. Если два сигнала неравновероятны, то оптимальный приемник преобразуется к виду, показанному на рис. 9.17[1]).

В данной главе рассматривался метод когерентного или синхронного приема. Такой прием выполняется согласованным фильтром или коррелятором. При рассмотрении амплитудной, частотной и фазовой манипуляций фазы принимаемых сигналов предполагались известными. Во многих случаях, однако, фаза несущей принимаемого сигнала неизвестна. Это может быть вызвано нестабильностью генератора передатчика и (или) гетеродина приемника, а также неизвестной длиной пути распространения. Таким образом, существует некоторая неопределенность фазы несущей принимаемого сигнала. При этих условиях когерентный прием (согласованный фильтр или корреляционный приемник) применить нельзя, поскольку в таких устройствах существенно точное знание прихода входного колебания. В таких случаях применяется некогерентный прием.

Некогерентный прием осуществляется путем подачи приходящего сигнала на детектор огибающей. Выходы детектора огибающей стробируются каждые T *сек* для принятия решения. Можно показать, что такой метод имеет худшую помехоустойчивость по сравнению с когерентным приемом, особенно при малых отношениях сигнал/шум.

ЗАДАЧИ

9.1. В двоичной системе связи одно из сообщений представляется прямоугольным импульсом $s(t)$, изображенным на рисунке. Другое сообщение передается отсутствием импульса. Импульсная реакция согласованного фильтра $h(t)=s(T-t)=s(t)$. Вычислите отношение мощности сигнала к мощности шума $s_0^2(t)/\overline{n_0^2}(t)$ при $t=T$. Предполагается, что шум белый с плотностью мощности $N/2$.



Решено использовать простейший RC-фильтр (см. рисунок) вместо согла-

[1]) См., например, Дж. Возенкрафт, И. Джекобс. Теоретические основы техники связи. М., «Мир», 1969. — *Прим. ред.*

сованного фильтра в приемнике. Определите максимальное отношение мощности сигнала к мощности шума $[s_o^2(t)/\overline{n_o^2(t)}]$, которое может быть получено при таком фильтре, и сравните его с тем, которое дает согласованный фильтр.

Указание. Обратите внимание на то, что сигнал $s_o(t)$ максимален при $t=T$. Отношение сигнал/шум является функцией постоянной времени RC. Найдите значение RC, которое дает максимум отношения сигнал/шум на выходе фильтра.

9.2. Найдите функцию передачи согласованного фильтра для гауссова (колокольного) импульса

$$s(t) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-t^2/2\sigma^2}.$$

Шум в канале — белый, со спектральной плотностью мощности $N/2$.

Определите максимальное отношение сигнал/шум на выходе этого фильтра.

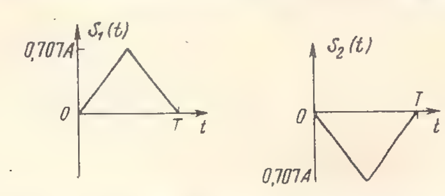
9.3. Покажите, что $s_o(t)$ — выходной сигнал фильтра, согласованного со входным сигналом $s(t)$, симметричен относительно $t=T$.

9.4. Два сообщения передаются посредством посылки и паузы при помощи импульса, показанного на рисунке.

Синтезируйте оптимальный приемник, если шум в канале является белым со спектральной плотностью мощности $N/2$ ($N=10^{-4}$ *вт/гц*).

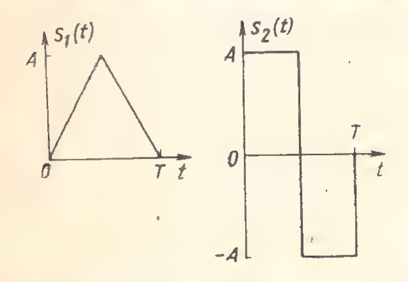
Определите вероятность ошибки оптимального приемника, предполагая, что вероятность посылки $s(t)$ равна 0,5.

9.5. При условии, что сообщения задачи 9.4 передаются двумя импульсами, изображенными на рисунке, синтезируйте оптимальный приемник и определите вероятность ошибки. Сравните полученную схему со схемой задачи 9.4.



9.6. Гауссов случайный сигнал имеет нулевое среднее значение и среднее значение квадрата σ_n^2. Определите вероятность наблюдения значений сигнала, превосходящих $10\sigma_n$.

9.7. Синтезируйте оптимальный приемник для случая белого шума в канале, если два сообщения передаются колебаниями $s_1(t)$ и $s_2(t)$, показанными на рисунке.



Определите вероятность ошибки оптимального приемника. Сравните эту схему с той, которая использовалась при приеме одиночного треугольного импульса (как в задаче 9.4) или двух треугольных импульсов (как в задаче 9.5). Как сравнить эту схему со схемой приема сигналов частотной манипуляции?

9.8. В гл. 9 согласованный фильтр был найден для случая белого шума. Развивая этот подход, получите согласованный фильтр для окрашенного шума (т. е. шума с неравномерной плотностью мощности) с заданной плотностью мощности $S_n(\omega)$.

Указание. Положите в неравенстве Шварца [ф-ла (9.7а)] $F_1(\omega) = \mathfrak{S}(\omega)H(\omega)$, $F_2(\omega) = S(\omega)/\mathfrak{S}(\omega)$, где $\mathfrak{S}(\omega)$ получается разложением $S_n(\omega)$ на множители: $S_n(\omega) = \mathfrak{S}(\omega)\mathfrak{S}(-\omega)$ и $\mathfrak{S}(\omega)$ имеет все полюсы и нули в левой полуплоскости комплексной частоты.

Литература

Bracewell R. M. The Fourier Transform and Its Applications, Mc Graw-Hill, New York, 1965.

Craig E. J. Laplace and Fourier Transforms for Electrical Engineers, Holt, Rinehart, and Winston, New York, 1964.

Javid M. and E. Brenner, Analysis, Transmission and Filtering of Signals, Mc Graw-Hill, New York, 1963.

Lathi B. P., Signals, Systems and Communication John Wiley and Sons. New York, 1965.

Marshall J. L., Signal Theory, International Textbook Co., Scranton, Pa.

Papoulis A., The Fourier Integral and Its Applications, Mc Graw—Hill, New York, 1962.

Black H. S., Modulation Theory, D. Van Nostrand Co., Princeton, N. J., 1953.

Bennett W. R. and J. R. Davey, Data Transmission, Mc Graw—Hill, New York, 1965.

Downing J. J., Modulation Systems and Noise, Prentice—Hall, Englewood Cliffs, N. J., 1964.

Freeman J. J., Principles of Noise, John Wiley and Sons, New York, 1958.

Hancock J., Principles of Communication Theory, Mc Graw—Hill, New York, 1961.

Panter P. F., Modulation, Noise and Spectral Analysis, Mc Graw—Hill, New York, 1965.

Rowe H. E., Signals and Noise in Communication Systems. D. Van Nostrand Co., Princeton, N. J., 1965.

Schwartz M., Information Transmission, Modulation and Noise, Mc Graw—Hill, New York, 1959.

Abramson N., Information Theory and Coding, Mc Graw—Hill, New York, 1963.

Harman W. W., Principles of the Statistical Theory of Communication, Mc Graw—Hill, New York, 1963.

Lathi B. P., An Introduction to Random Signals and Communication Theory, International Textbook Co., 1968.

Reza F. M., An Introduction to Information Theory, Mc Graw—Hill, New York, 1961.

Schwartz M., W. R. Bennett and S. Stein, Communication Systems and Techniques, Mc Graw—Hill, New York, 1966.

Wozencraft J. M., and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, New York, 1965.

СОДЕРЖАНИЕ

Литература

Bracewell R. M. The Fourier Transform and Its Applications, Mc Graw-Hill, New York, 1965.
Craig E. J. Laplace and Fourier Transforms for Electrical Engineers, Holt, Rinehart, and Winston, New York, 1964.
Javid M. and E. Brenner, Analysis, Transmission and Filtering of Signals, Mc Graw-Hill, New York, 1963.
Lathi B. P., Signals, Systems and Communication John Wiley and Sons. New York, 1965.
Marshall J. L., Signal Theory, International Textbook Co., Scranton, Pa.
Papoulis A., The Fourier Integral and Its Applications, Mc Graw—Hill, New York, 1962.
Black H. S., Modulation Theory, D. Van Nostrand Co., Princeton, N. J., 1953.
Bennett W. R. and J. R. Davey, Data Transmission, Mc Graw—Hill, New York, 1965.
Downing J. J., Modulation Systems and Noise, Prentice—Hall, Englewood Cliffs, N. J., 1964.
Freeman J. J., Principles of Noise, John Wiley and Sons, New York, 1958.
Hancock J., Principles of Communication Theory, Mc Graw—Hill, New York, 1961.
Panter P. F., Modulation, Noise and Spectral Analysis, Mc Graw—Hill, New York, 1965.
Rowe H. E., Signals and Noise in Communication Systems. D. Van Nostrand Co., Princeton, N. J., 1965.
Schwartz M., Information Transmission, Modulation and Noise, Mc Graw—Hill, New York, 1959.
Abramson N., Information Theory and Coding, Mc Graw—Hill, New York, 1963.
Harman W. W., Principles of the Statistical Theory of Communication, Mc Graw—Hill, New York, 1963.
Lathi B. P., An Introduction to Random Signals and Communication Theory, International Textbook Co., 1968.
Reza F. M., An Introduction to Information Theory, Mc Graw—Hill, New York, 1961.
Schwartz M., W. R. Bennett and S. Stein, Communication Systems and Techniques, Mc Graw—Hill, New York, 1966.
Wozencraft J. M., and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, New York, 1965.

СОДЕРЖАНИЕ

СИСТЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ